Observational bias correction in data assimilation and an overview of satellite data monitoring

Niels Bormann,
Hans Hersbach, Dick Dee, Mohamed Dahoui

NWP SAF Training Course on the Use of Satellite Data

ECMWF

Outline of part I: Observational bias correction

1. **Introduction**
   - Biases in **models**, **observations**, and **observation operators**
2. **Variational analysis and correction of observation bias**
   - The need for an **adaptive** system
   - Variational bias correction (**VarBC**)
3. **Limitations of VarBC and how to address them**
   - Interaction with **model bias**
   - Adding further **constraints**
4. **Summary**

Outline of part I: Observational bias correction

1. **Introduction**
   - Biases in **models**, **observations**, and **observation operators**
2. **Variational analysis and correction of observation bias**
   - The need for an **adaptive** system
   - Variational bias correction (**VarBC**)
3. **Limitations of VarBC and how to address them**
   - Interaction with **model bias**
   - Adding further **constraints**
4. **Summary**

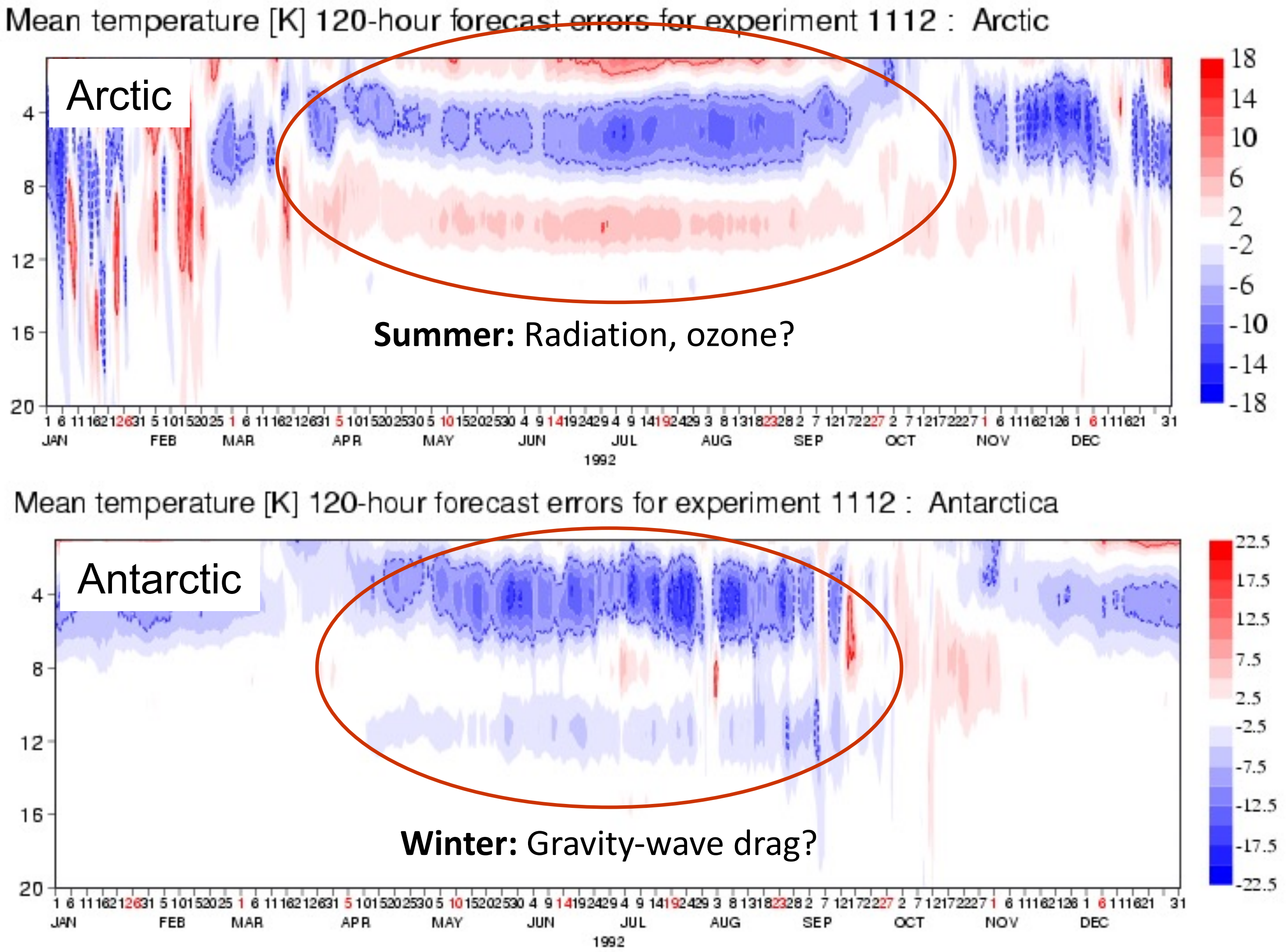
Biases are everywhere – in models, observations, observation operators

Example of a <u>model bias</u>:

Seasonal variations in temperature biases in the upper-stratosphere (T255L60 model used for the *ERA-Interim* reanalysis)

0.1hPa (65km)

40hPa (22km)

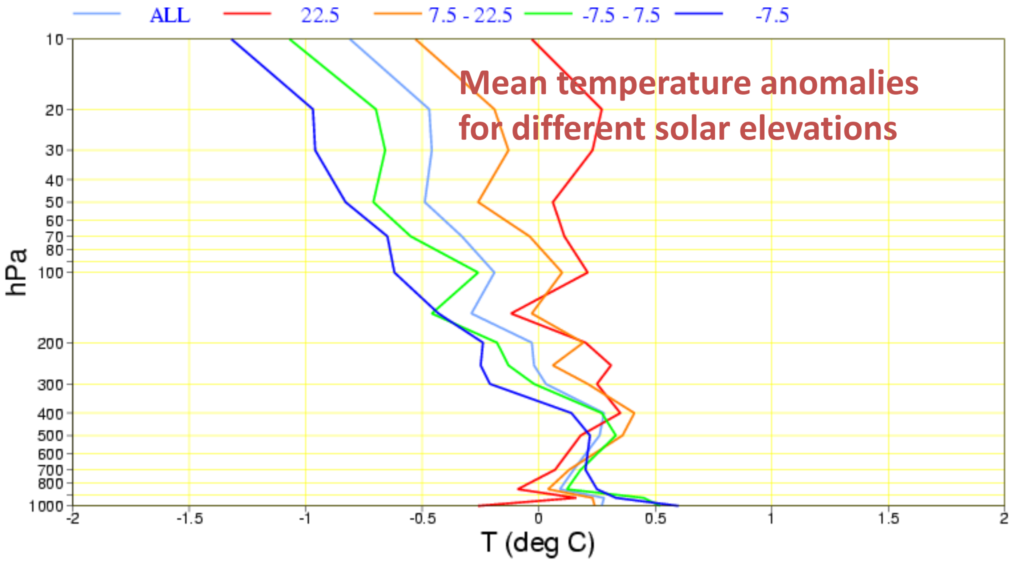


0.1hPa (65km)

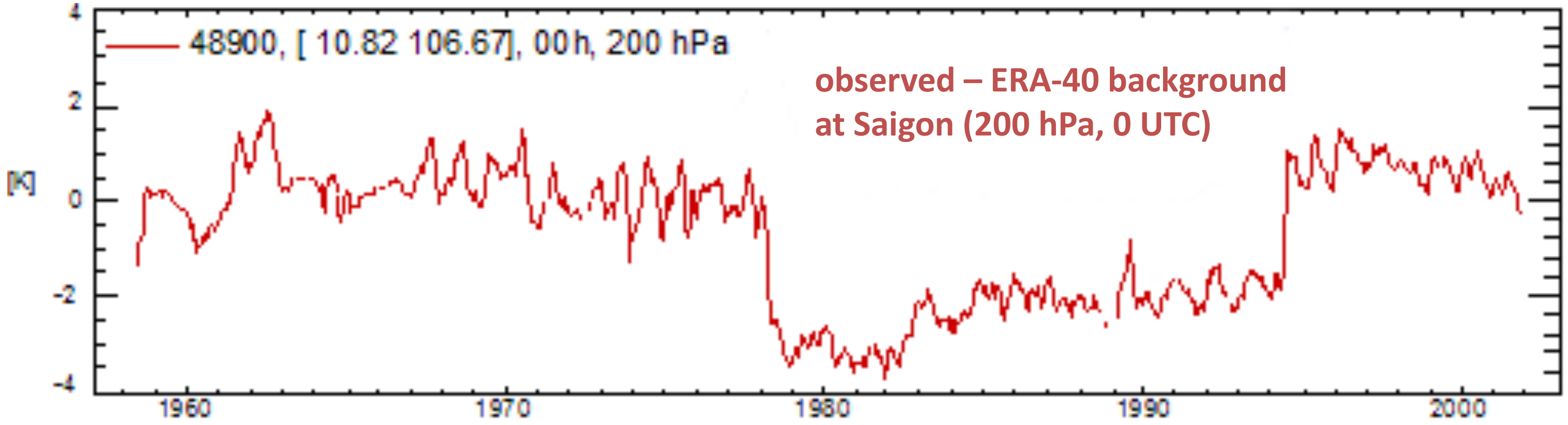
40hPa (22km)

Observation bias
E.g., : Radiosonde temperature observations

Daytime warm bias due to radiative heating of the temperature sensor
(depends on *solar elevation* and *equipment type*)



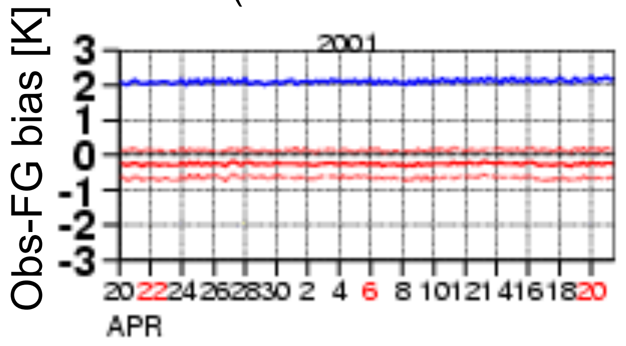
Bias changes due to change of equipment



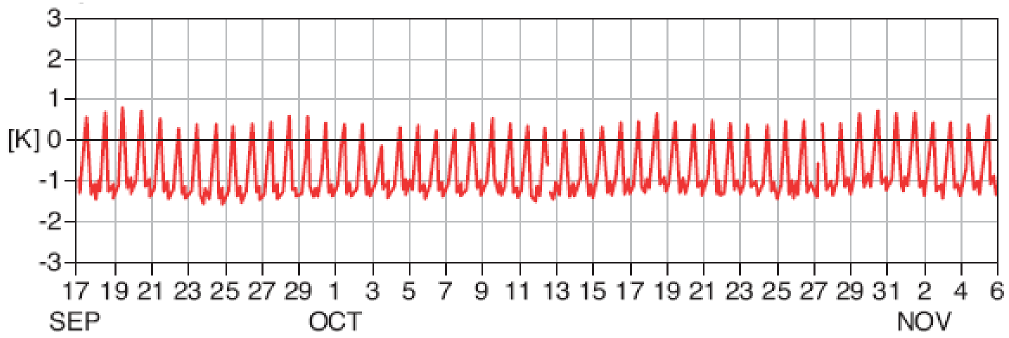
Observation and observation operator bias: Satellite radiances

Monitoring the background departures o-b (averaged in time and/or space):

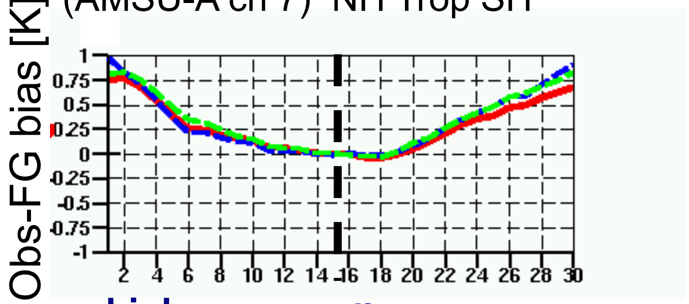
Constant bias (NOAA-14 HIRS channel 5)



Diurnal bias variation in a geostationary satellite

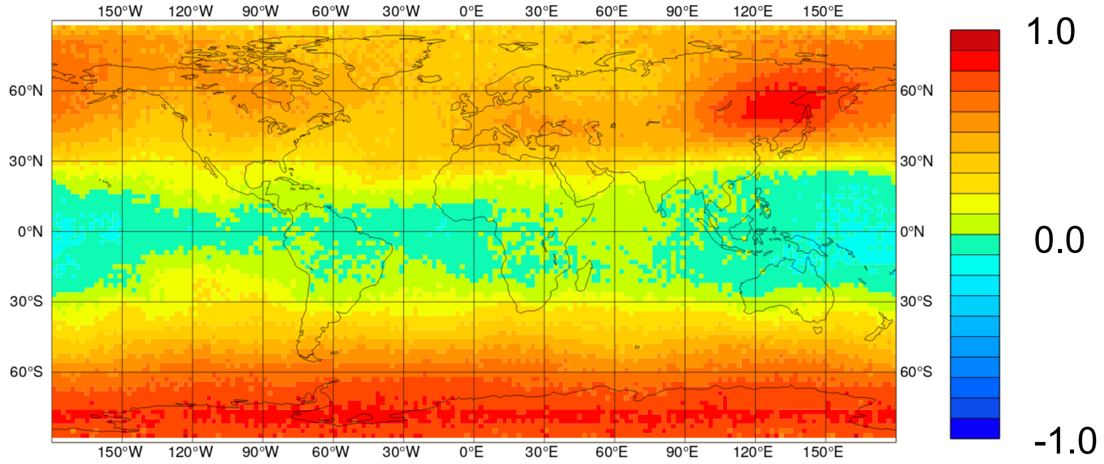


Bias depending on scan position (AMSU-A ch 7) NH Trop SH



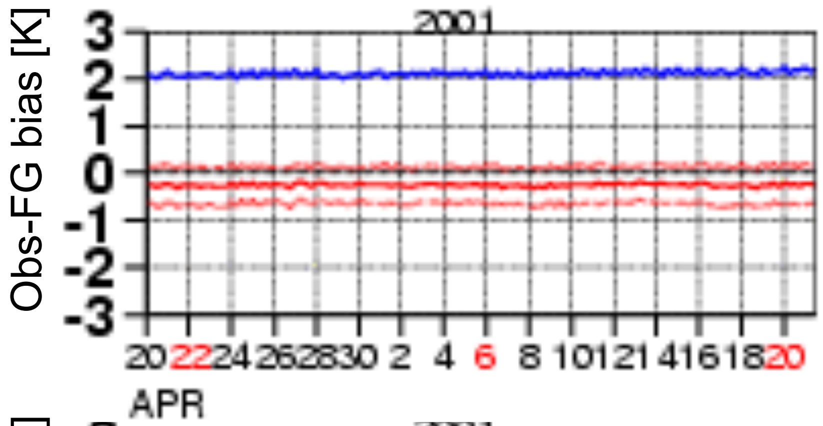
high zenith angle **nadir** **high zenith angle**

Air-mass dependent bias (AMSU-A ch 8)

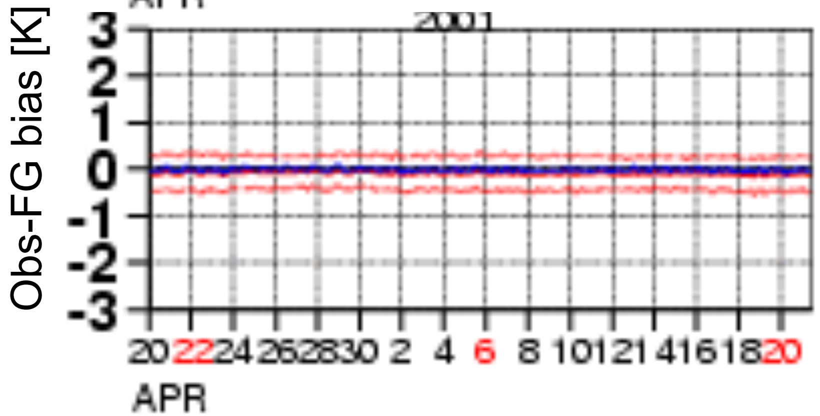


Observation and observation operator bias:
Satellite radiances – identifying sources of bias

Monitoring the background departures o-b (averaged in time and/or space):



HIRS channel 5 (peaking around 600hPa) on NOAA-14 satellite has +2.0K radiance bias against the background.

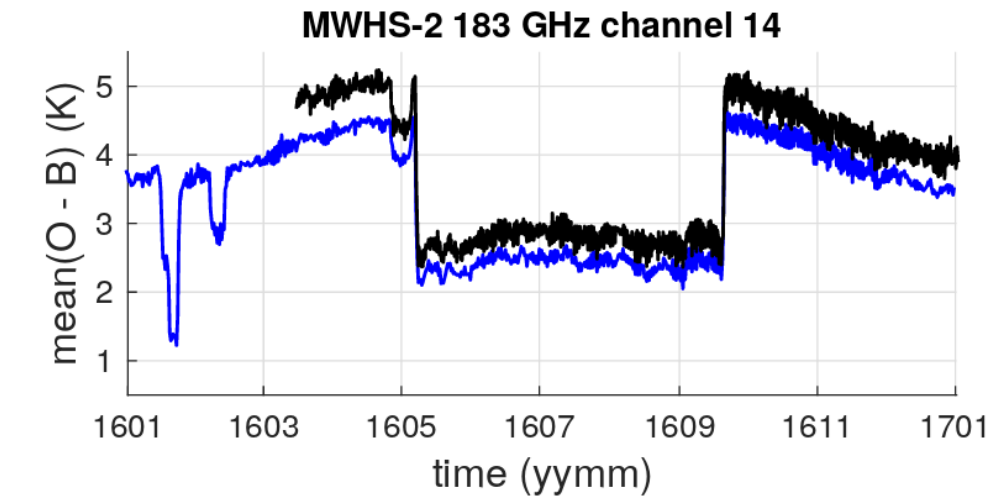


The same channel on the NOAA-16 satellite and other similar radiances have no bias against the background.

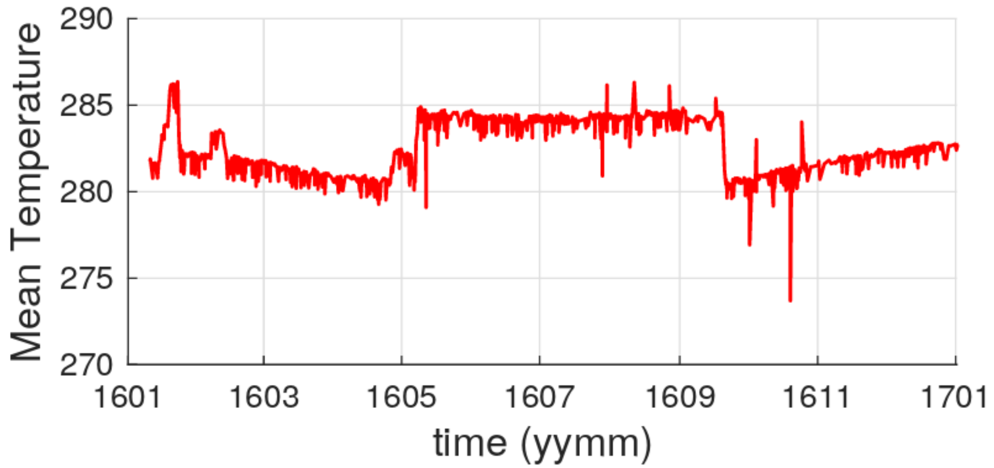
→ NOAA-14 channel 5 has an instrument bias (subsequently related to spectral response function for this channel).

Observation and observation operator bias: Satellite radiances – identifying sources of bias

A time-varying bias:



Similar bias changes in two NWP systems.



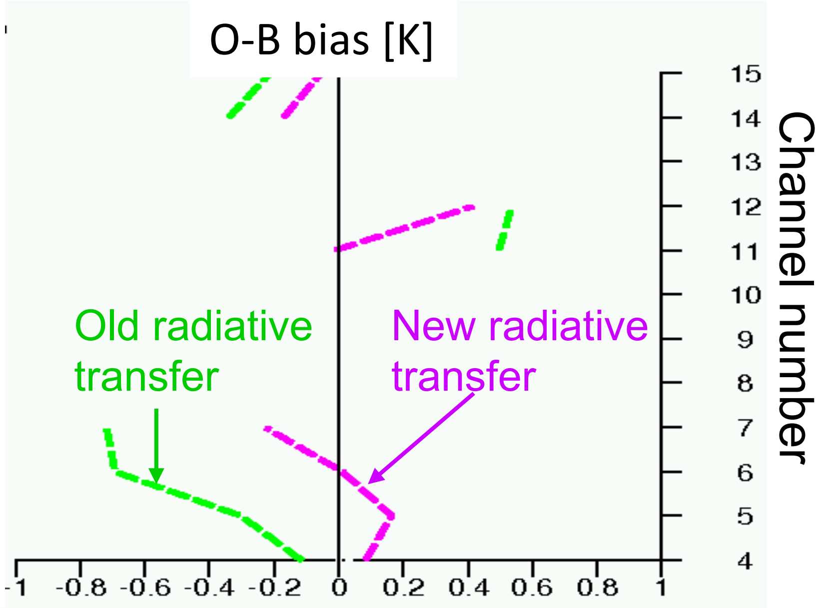
Bias changes apparently linked to the temperature of the instrument.

→ Channel affected by an instrument bias.

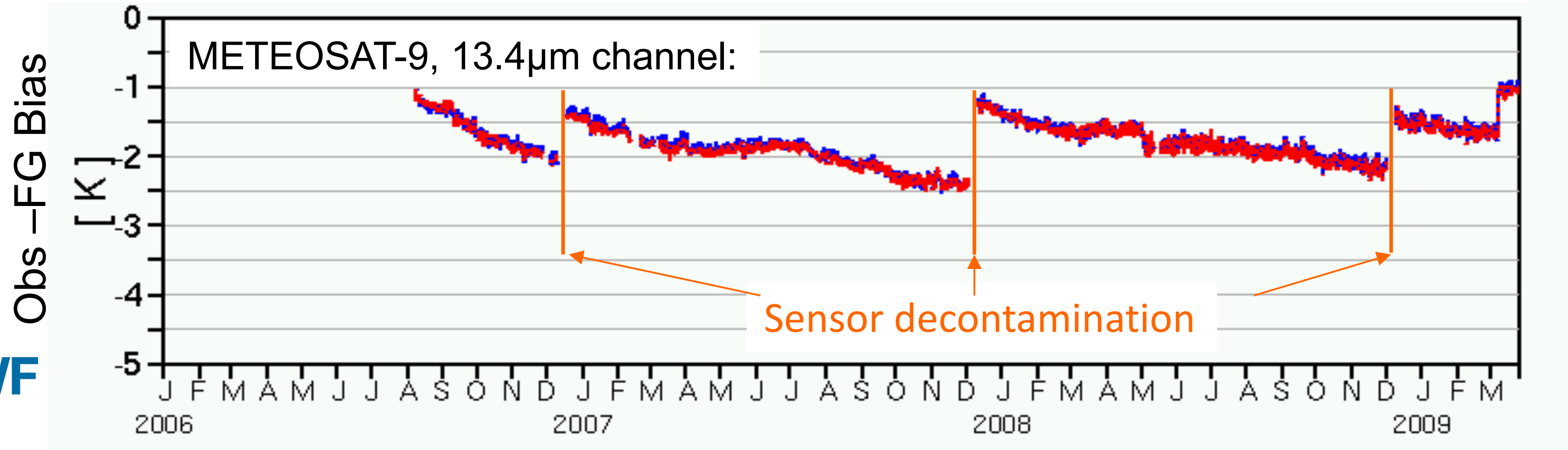
Observation and observation operator bias: Radiative transfer bias for satellite radiances

Examples of causes for biases in radiative transfer:
Bias in assumed concentrations of atmospheric gases
(e.g., CO_2, aerosols)
Biases in the spectroscopy
Neglected effects (e.g., clouds)
Incorrect spectral response function
…

Change in bias for HIRS resulting from an update of the Radiative Transfer model:



Drift in bias due to ice-build up on sensor, altering the spectral response of the channel:



Outline of part I: Observational bias correction

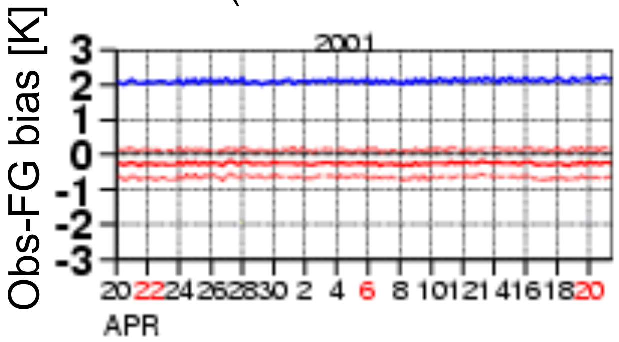
1. **Introduction**
   – Biases in **models**, **observations**, and **observation operators**
2. **Variational analysis and correction of observation bias**
   – The need for an **adaptive** system
   – Variational bias correction (**VarBC**)
3. **Limitations of VarBC and how to address them**
   – Interaction with **model bias**
   – Adding further **constraints**
4. **Summary**

How to address systematic errors?
The need for an adequate bias model

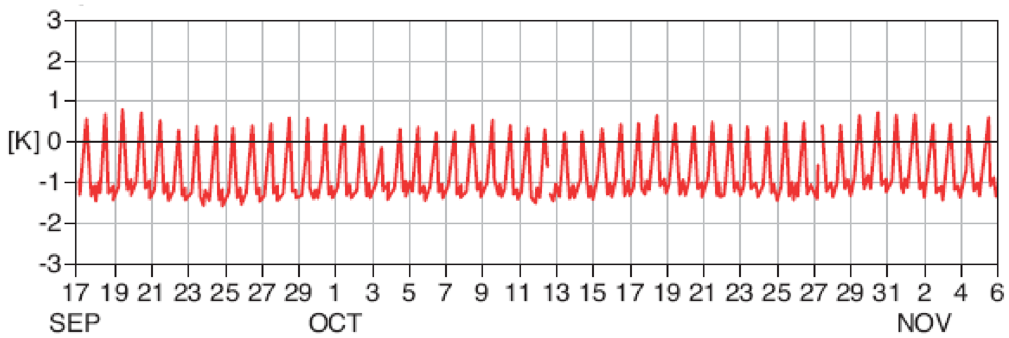
Prerequisite for any bias correction is a model for the bias (b(x,β)):

- Ideally, guided by the physical origins of the bias.
- In practice, bias models are derived empirically from observation monitoring after careful diagnosis of the bias.

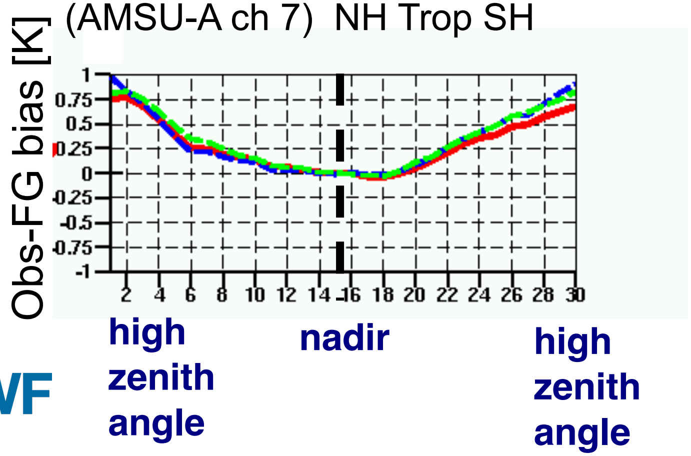
Constant bias (NOAA-14 HIRS channel 5)



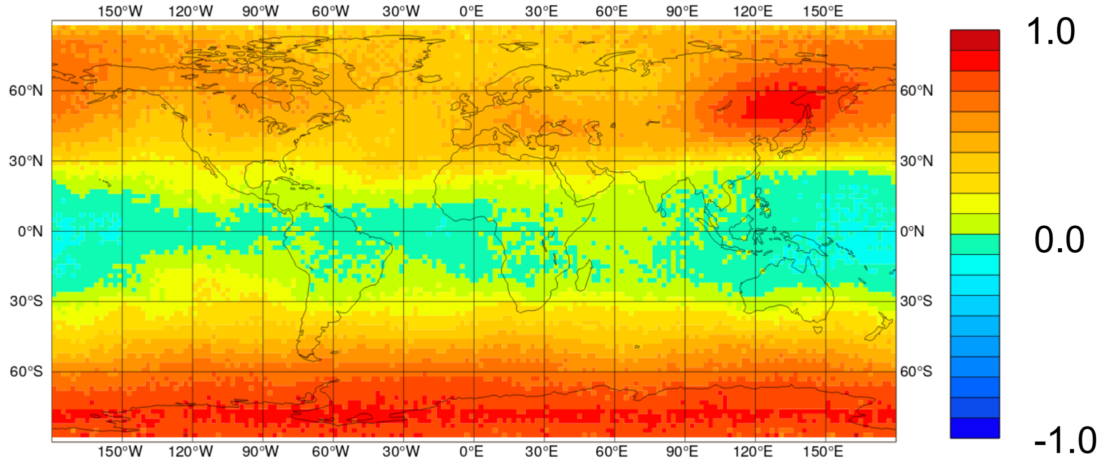
Diurnal bias variation in a geostationary satellite



Bias depending on scan position (AMSU-A ch 7) NH Trop SH



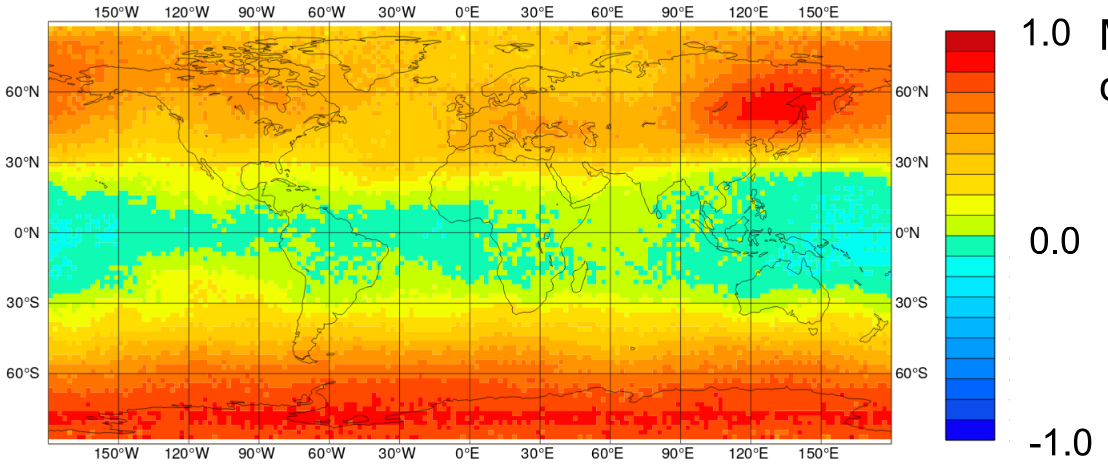
Air-mass dependent bias (AMSU-A ch 8)



How to address systematic errors? The need for an adequate bias model

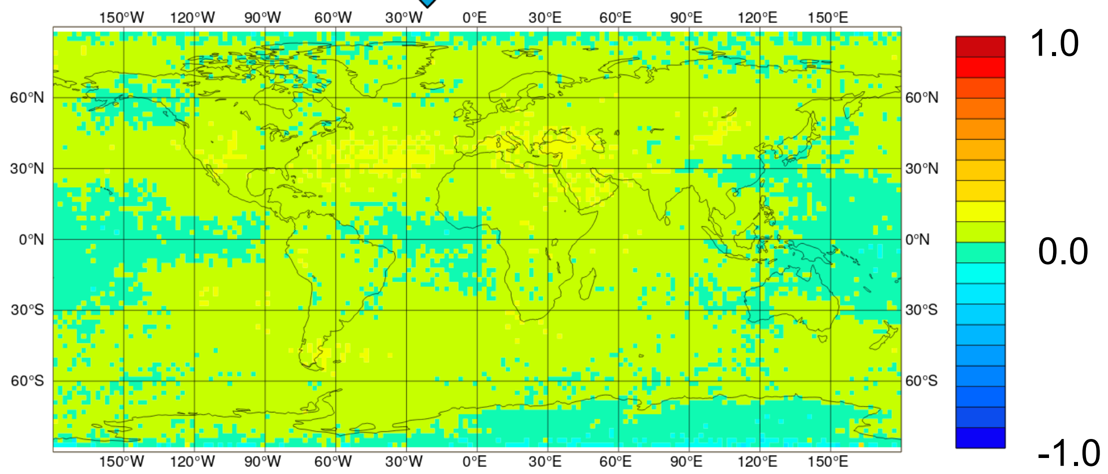
Prerequisite for any bias correction is a model for the bias ($b(x,\beta)$):

- For instance, a linear model with some predictors p_1, p_2, … p_n, and free parameters β_0, β_1, β_2, … β_n ("bias coefficients"): $b(\mathbf{x},\boldsymbol{\beta}) = \beta_0 + \beta_1 p_1 + \beta_2 p_2 + \ldots + \beta_n p_n$



Mean o-b before bias correction

After bias correction



The example uses a linear bias model with a constant β_0 and four layer thicknesses as predictors (1000-300hPa, 200-50hPa, 50-5hPa,10-1hPa thickness) + a model for scan-bias

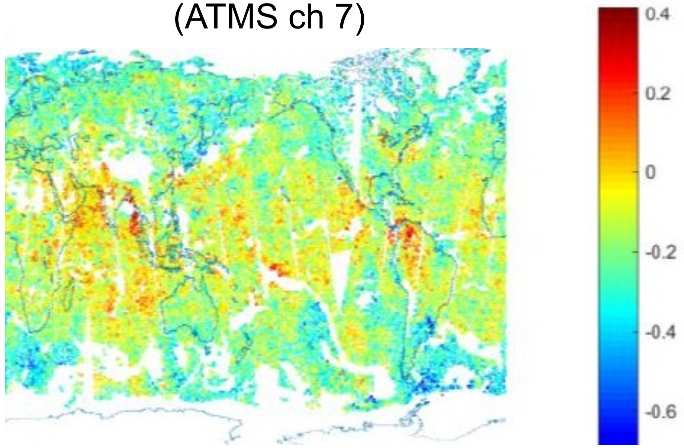
ECMWF

How to address systematic errors?
The need for an adequate bias model

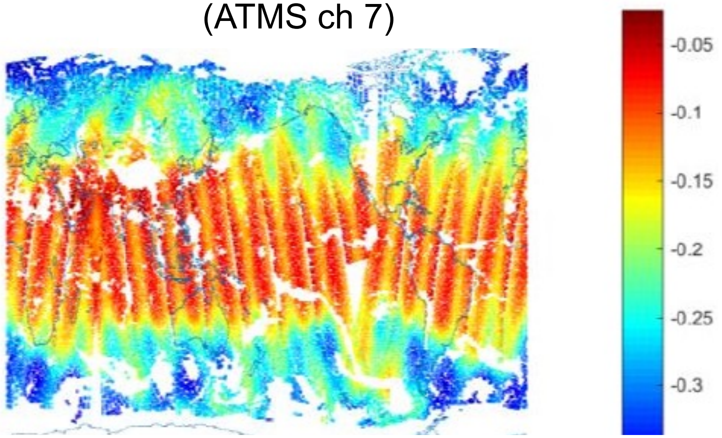
Prerequisite for any bias correction is a model for the bias (b(x,β)):

- For instance, a linear model with some predictors p_1, p_2, … p_n, and free parameters β_0, β_1, β_2, … β_n ("bias coefficients"): $b(\mathbf{x},\boldsymbol{\beta}) = \beta_0 + \beta_1 p_1 + \beta_2 p_2 + \ldots + \beta_n p_n$

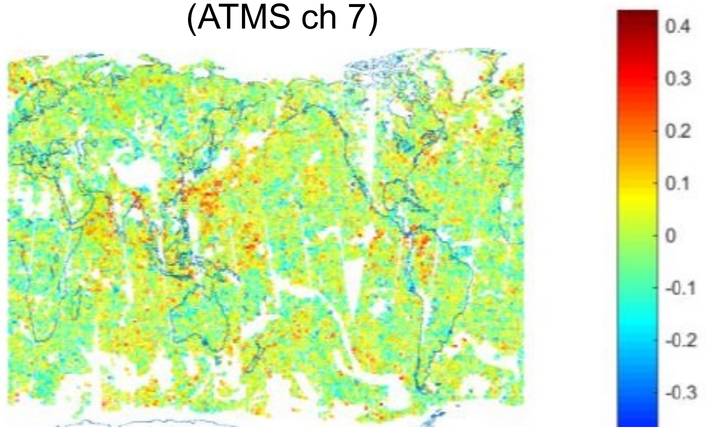
Uncorrected departure [K]
(ATMS ch 7)



Bias correction [K]
(ATMS ch 7)

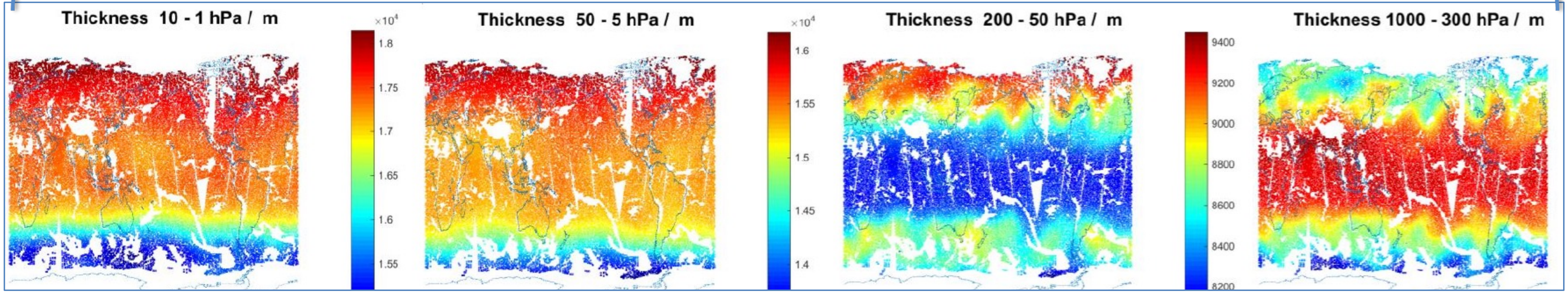


Corrected departure [K]
(ATMS ch 7)



+ offset + model for scan-bias

Airmass predictors p_i



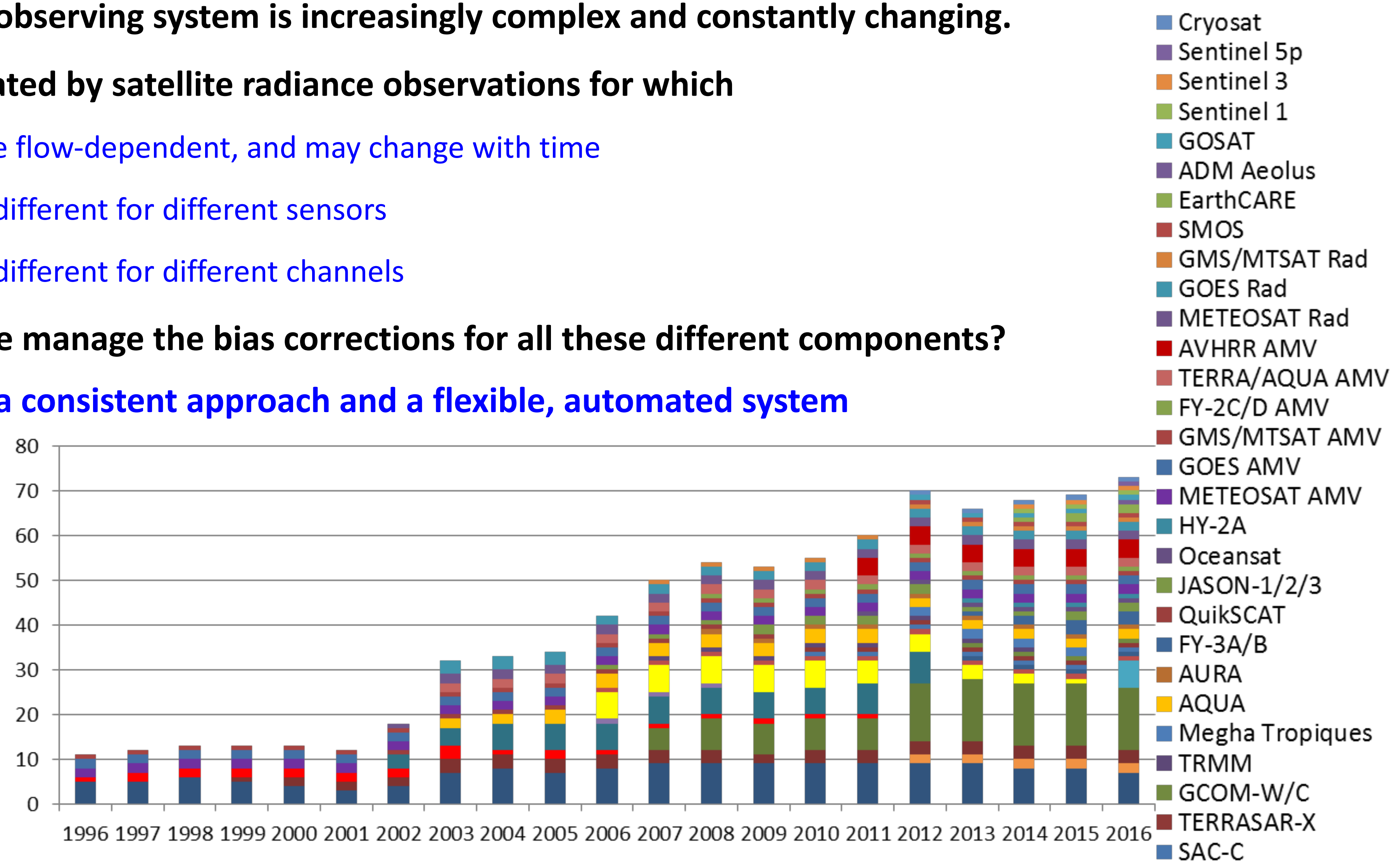
Offline bias correction
(as used for satellite radiances at ECMWF before 2006)

- **Bias coefficients** were estimated **off-line** for each satellite/sensor/channel from past background departures, and stored in files **(Harris and Kelly 2001).**
  - Using a regression procedure.
  - Typically based on 2 weeks of background departures.
  - After careful masking and data selection

- Bias coefficients were then applied to new data and kept fixed until an update was considered necessary.

The need for an adaptive bias correction system

- **The global observing system is increasingly complex and constantly changing.**
- **It is dominated by satellite radiance observations for which**
  - biases are flow-dependent, and may change with time
  - they are different for different sensors
  - they are different for different channels
- **How can we manage the bias corrections for all these different components?**
  - **Requires a consistent approach and a flexible, automated system**



Variational bias correction: General Idea

The **bias** in a given instrument/channel is described by (a few) **bias parameters**:
typically, these are functions of air-mass and scan-position (the **predictors**)
These parameters can be estimated in a variational analysis along with the model state (Derber and Wu, 1998 at NCEP, USA)

The original problem:

$$J(\mathbf{x}) = \overbrace{(\mathbf{x}_b - \mathbf{x})^T \mathbf{B}^{-1} (\mathbf{x}_b - \mathbf{x})}^{J_b\text{: background constraint}} + \underbrace{[\mathbf{y} - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - \mathbf{h}(\mathbf{x})]}_{J_o\text{: observation constraint}}$$

The modified problem:

$$\begin{aligned} J(\mathbf{x},\boldsymbol{\beta}) = & \overbrace{(\mathbf{x}_b - \mathbf{x})^T \mathbf{B}_x^{-1} (\mathbf{x}_b - \mathbf{x})}^{J_b\text{: background constraint for } \mathbf{x}} + \overbrace{(\boldsymbol{\beta}_b - \boldsymbol{\beta})^T \mathbf{B}_\beta^{-1} (\boldsymbol{\beta}_b - \boldsymbol{\beta})}^{J_\beta\text{: background constraint for } \beta} \\ & + \underbrace{[\mathbf{y} - \mathbf{b}_o(\mathbf{x},\boldsymbol{\beta}) - \mathbf{h}(\mathbf{x})]^T \mathbf{R}^{-1} [\mathbf{y} - \mathbf{b}_o(\mathbf{x},\boldsymbol{\beta}) - \mathbf{h}(\mathbf{x})]}_{J_o\text{: bias-corrected observation constraint}} \end{aligned}$$

Parameter estimates from previous analysis

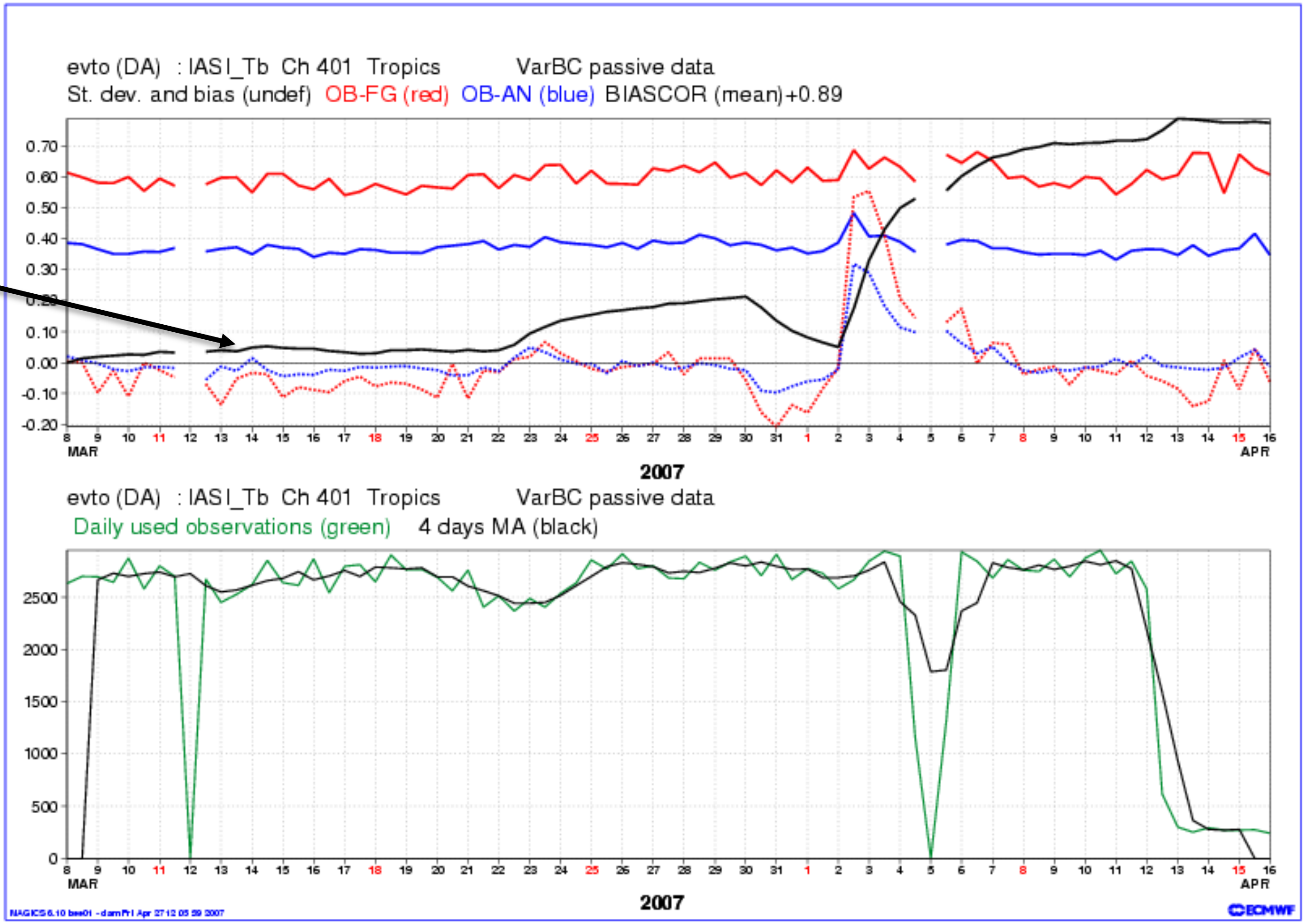
A model for the observation bias

Example of using VarBC (I):

Spinning up a new instrument – IASI on MetOp A

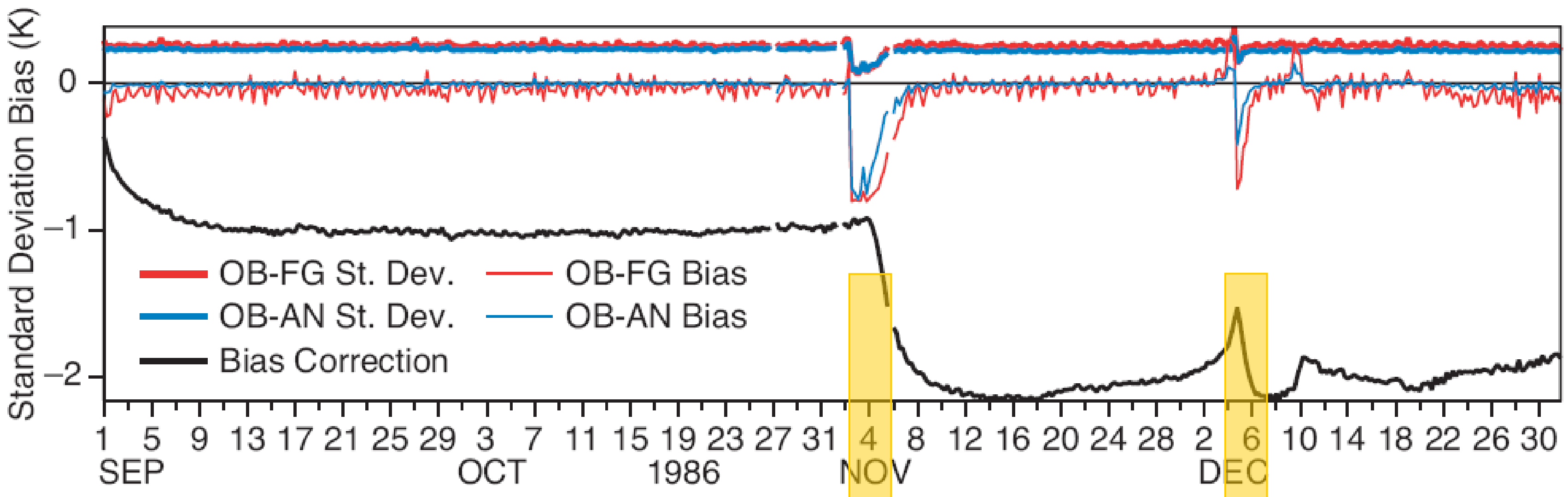
- IASI is an interferometer with 8461 channels
- Initially unstable – data gaps, preprocessing changes

Bias correction
+ 0.89

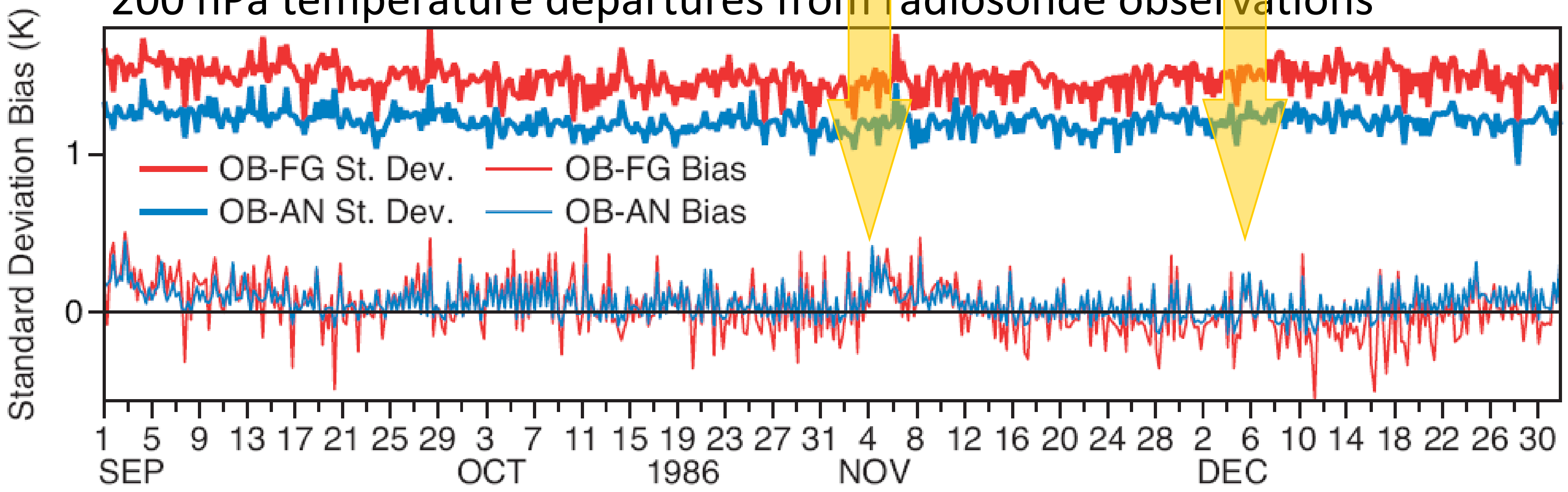


Example of using VarBC (II):

Reaction of NOAA-9 MSU channel 3 bias corrections following a cosmic storm



200 hPa temperature departures from radiosonde observations



ECMWF

Current observational bias correction at ECMWF

Observations treated by VarBC in the operational ECMWF system:

- Radiances
- Ozone
- Aircraft data
- Ground-based radar precipitation

Other automated bias corrections, but outside 4D-Var:

- Surface pressure
- Radiosonde temperature and humidity
- Soil moisture (in SEKF surface analysis)

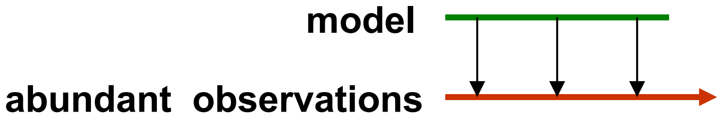
Outline of part I: Observational bias correction

1. **Introduction**
   - Biases in **models**, **observations**, and **observation operators**
2. **Variational analysis and correction of observation bias**
   - The need for an **adaptive** system
   - Variational bias correction (**VarBC**)
3. **Limitations of VarBC and how to address them**
   - Interaction with **model bias**
   - Adding further **constraints**
4. **Summary**

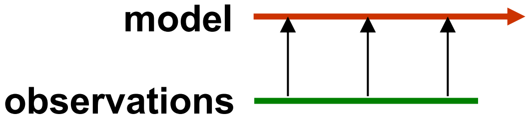
Limitations of VarBC:
Interaction with model bias

VarBC introduces extra degrees of freedom in the variational analysis, to help improve the fit to the (bias-corrected) observations.

It works well (even if the model is biased) when the analysis is strongly constrained by observations:



It does not work as well when there are large model biases and observation biases are poorly constrained (e.g., few anchoring observations; many bias-corrected observations with similar characteristics):



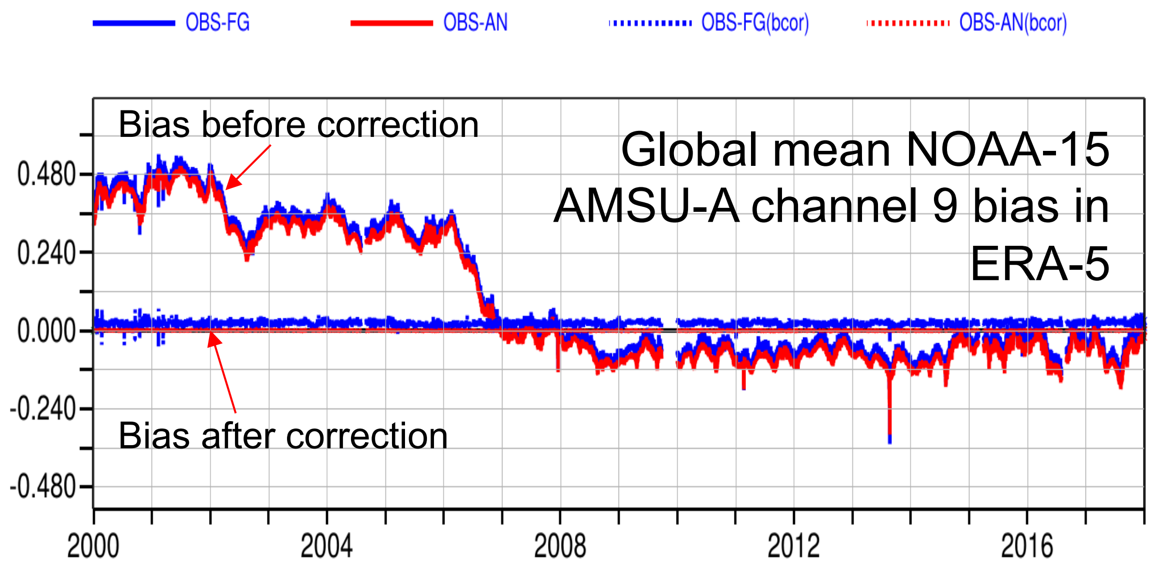
VarBC is not designed to correct model biases: Need different methods to estimate mode error (e.g., **weak-constraint 4D-Var**).

Limitations of VarBC: Interaction with model bias and the role of anchor observations

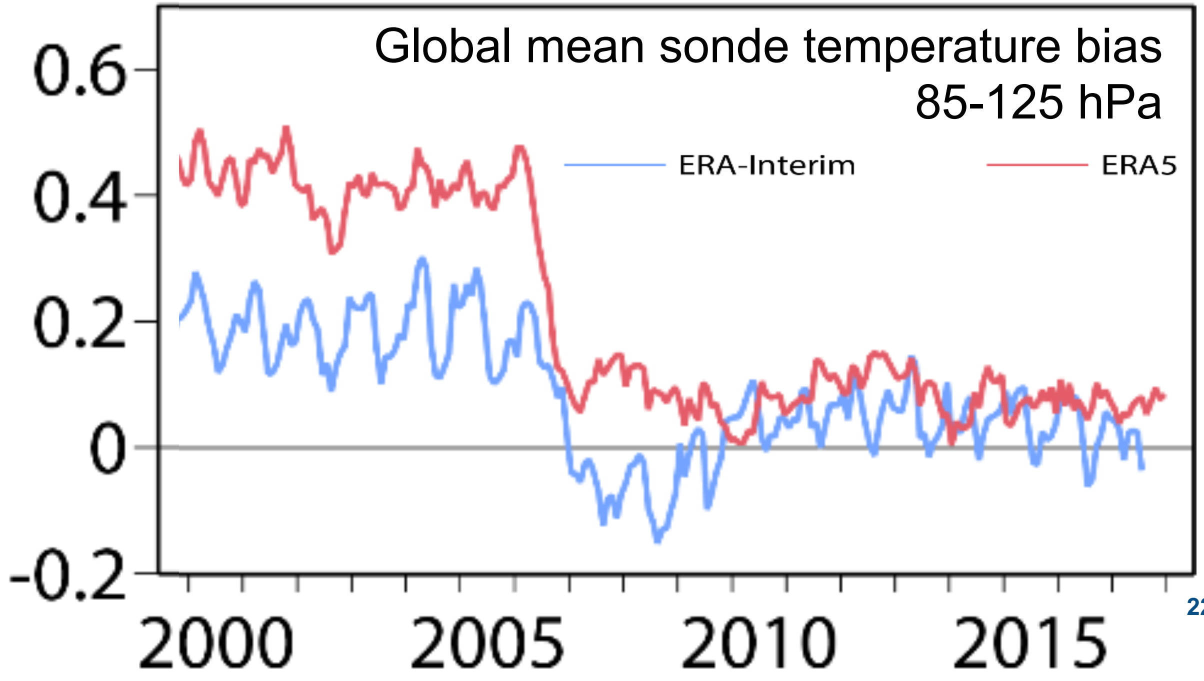
Example: Stratospheric temperature biases

- Model biases affect the bias correction in the absence of sufficient anchor observations.
- GNSS-RO provides a good anchor from mid-2006.

- The solution of the bias correction is also affected by other aspects, including the background error covariance.



Increased availability of GNSS-RO data



ECMWF

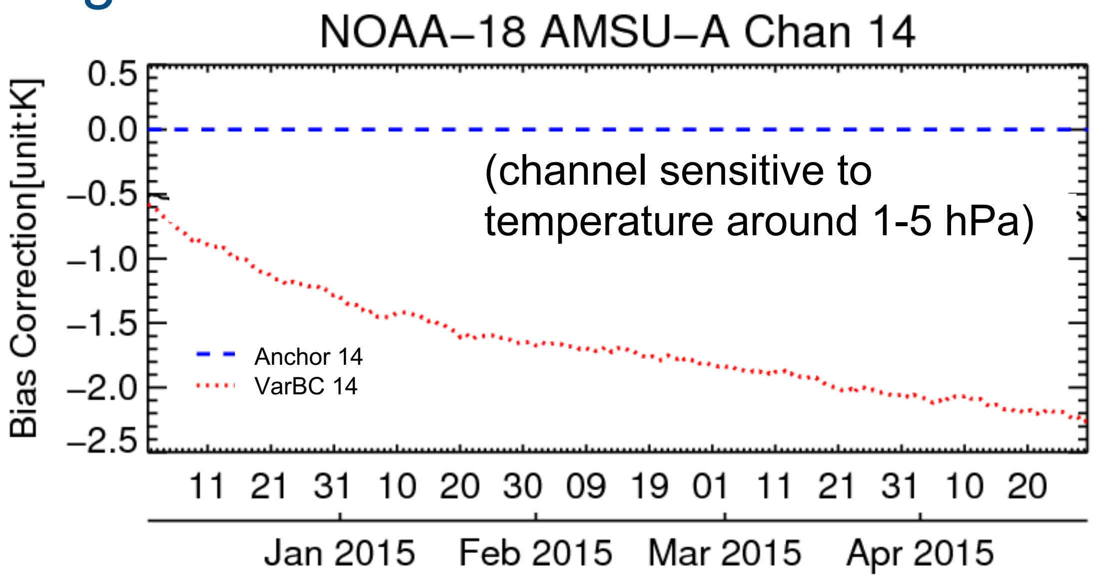
Limitations of VarBC:
Interaction with model bias - selecting an anchor observation

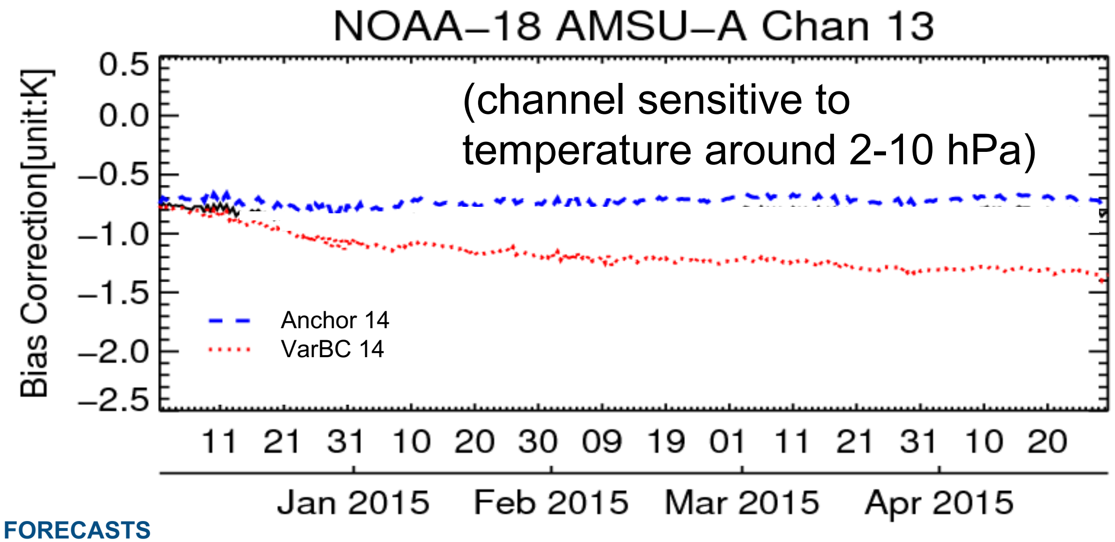
Example: Upper stratospheric temperature biases

- Unrealistic drift in the bias corrections due to model bias (red line)
- Additional ***anchoring*** can be imposed through assimilating AMSU-A channel 14 without a bias correction (blue line)

- Other anchoring in the ECMWF system: selected ozone-sensitive IR channels

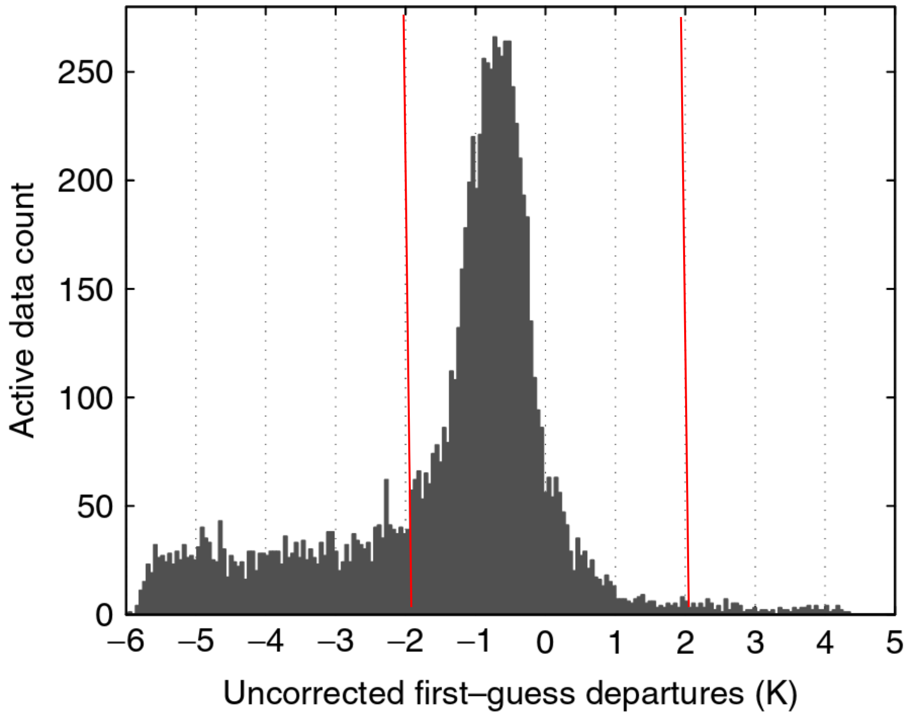
- Other ways to penalize (too) large bias corrections: Constrained VarBC (Han and Bormann 2016)





Limitations of VarBC:
Other pit-falls: Removing the signal

- **Avoid** bias correction models with **too many predictors**, to avoid correcting for situation-dependent background errors/biases to be incorrectly removed.
- Beware of interaction between VarBC and **departure-based quality control** and asymmetric distributions:
    - Can lead to unwanted drifts in the population after QC



Histogram of IR window channel departures with cold cloud tail.

Summary of part I: Observational bias correction

- **Biases are everywhere:**
  - Most observations cannot be usefully assimilated without bias adjustments.

- **Manual estimation of biases in satellite data is practically impossible.**
- **Bias estimates can be updated automatically during data assimilation.**

- **Variational bias correction works best in situations where:**
  - there is sufficient redundancy in the data; or
  - there are no large model biases

Challenges:

- How to develop good bias models for observations.
  - Potential for machine learning?
- How to separate observation bias from model bias.

Part II: Satellite data monitoring

ECMWF satellite data monitoring

https://www.ecmwf.int/en/forecasts/quality-our-forecasts/monitoring-observing-system#Satellite

ECMWF

En Fr De Search site... Contact Niels Bormann

Home | About | Forecasts | Computing | Research | Learning | Library

Charts | Datasets | **Quality of our forecasts** | Documentation and support | Accessing forecasts

The EUMETSAT Network of Satellite Application Facilities

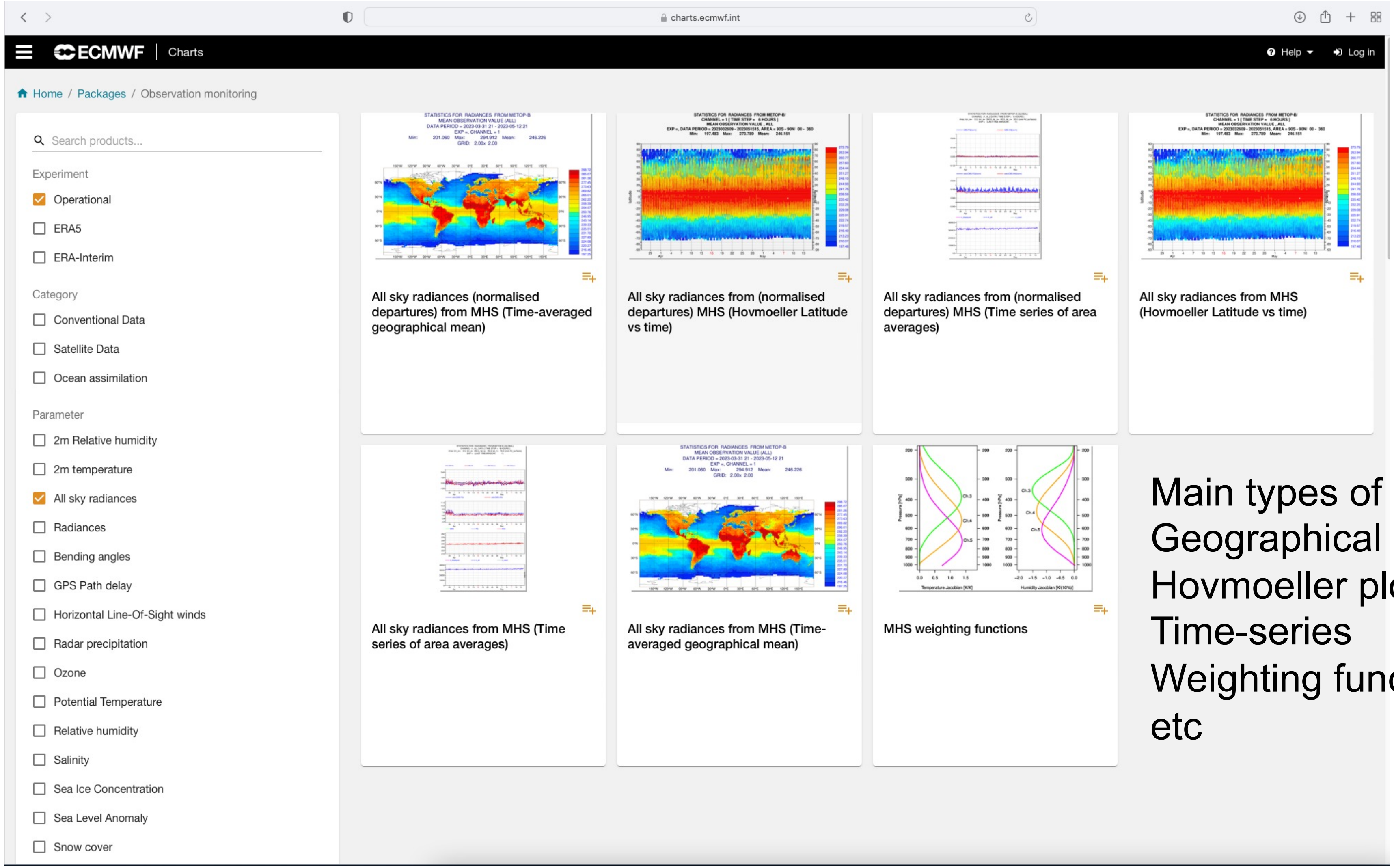
NWP SAF
Numerical Weather Prediction

The satellite monitoring is organised by satellite data types:

- All sky Microwave radiances
- Clear sky Microwave radiances
- Infrared sounding radiances
- GPS Radio Occultation (GPSRO)
- Atmospheric Motion Vectors
- Geostationary radiances
- Surface wind
- Soil Moisture and Ocean Salinity (SMOS)
- NESDIS Snow and Ice Mapping System (IMS)
- Ozone monitoring
- Soil moisture
- Significant wave height

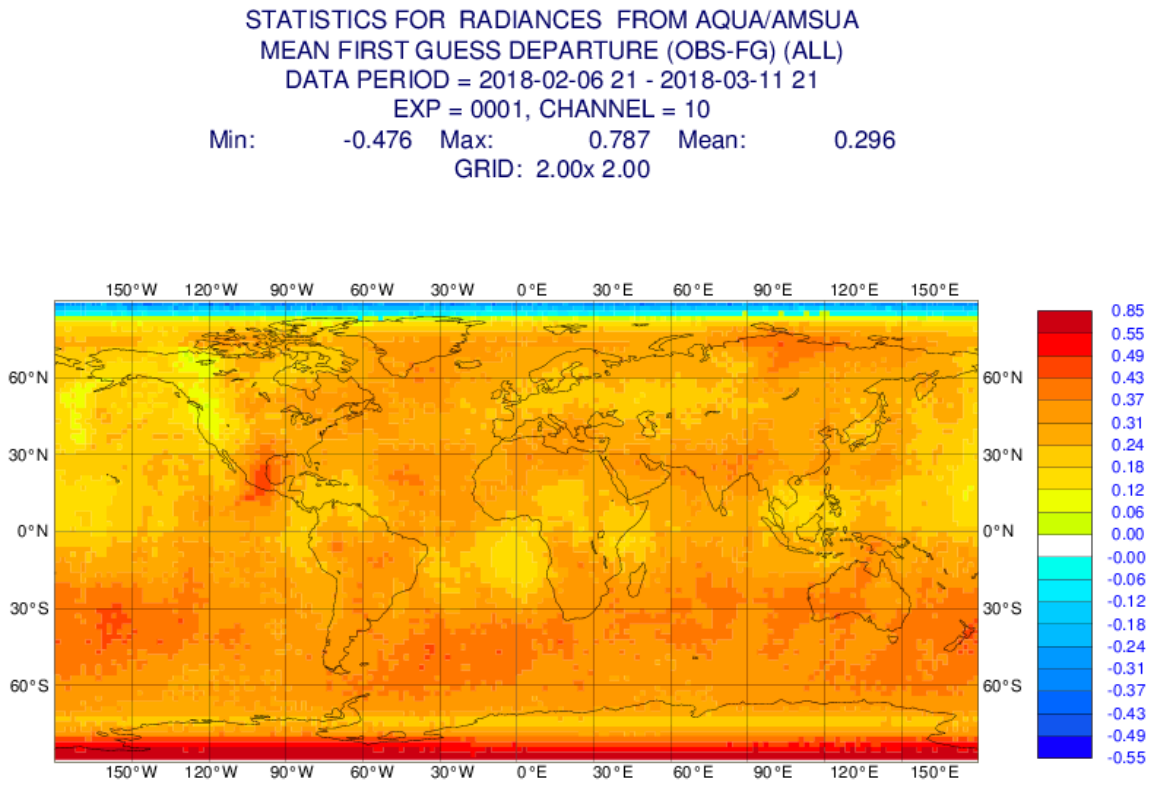
ECMWF satellite data monitoring

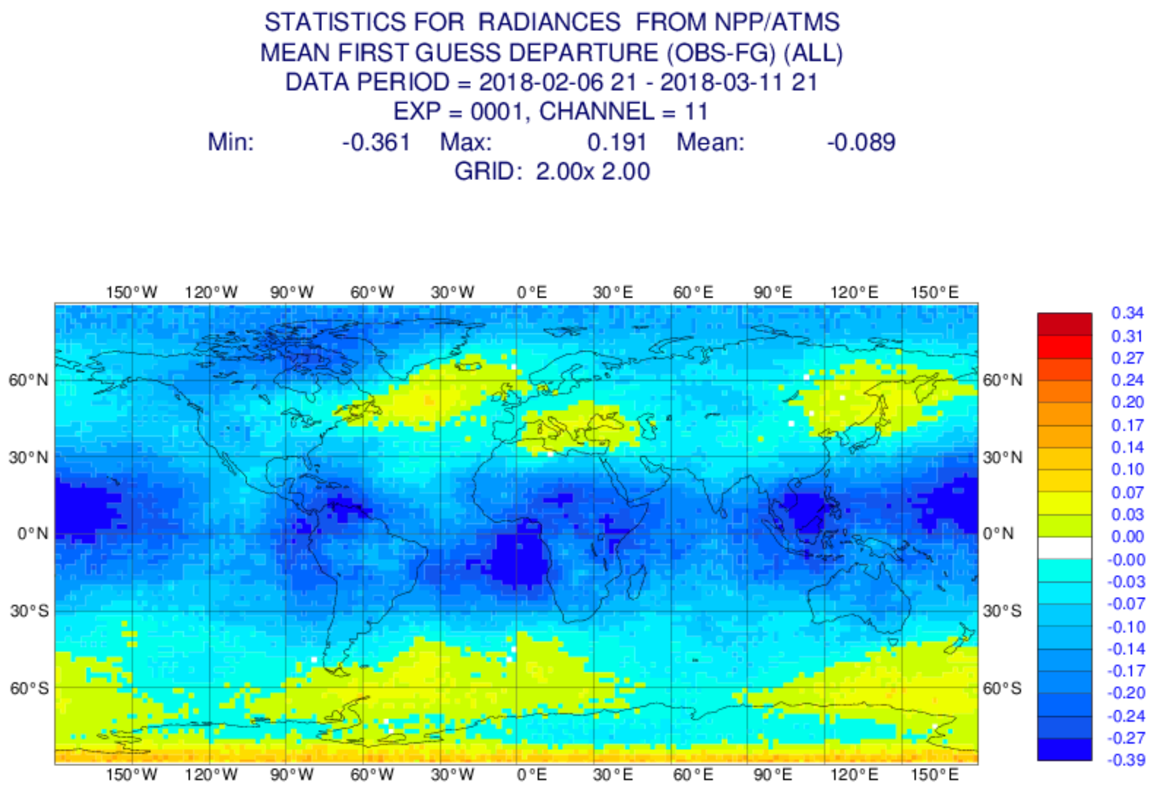
https://www.ecmwf.int/en/forecasts/quality-our-forecasts/monitoring-observing-system#Satellite



Main types of plots available:
Geographical maps
Hovmoeller plots
Time-series
Weighting functions (for some instruments)
etc

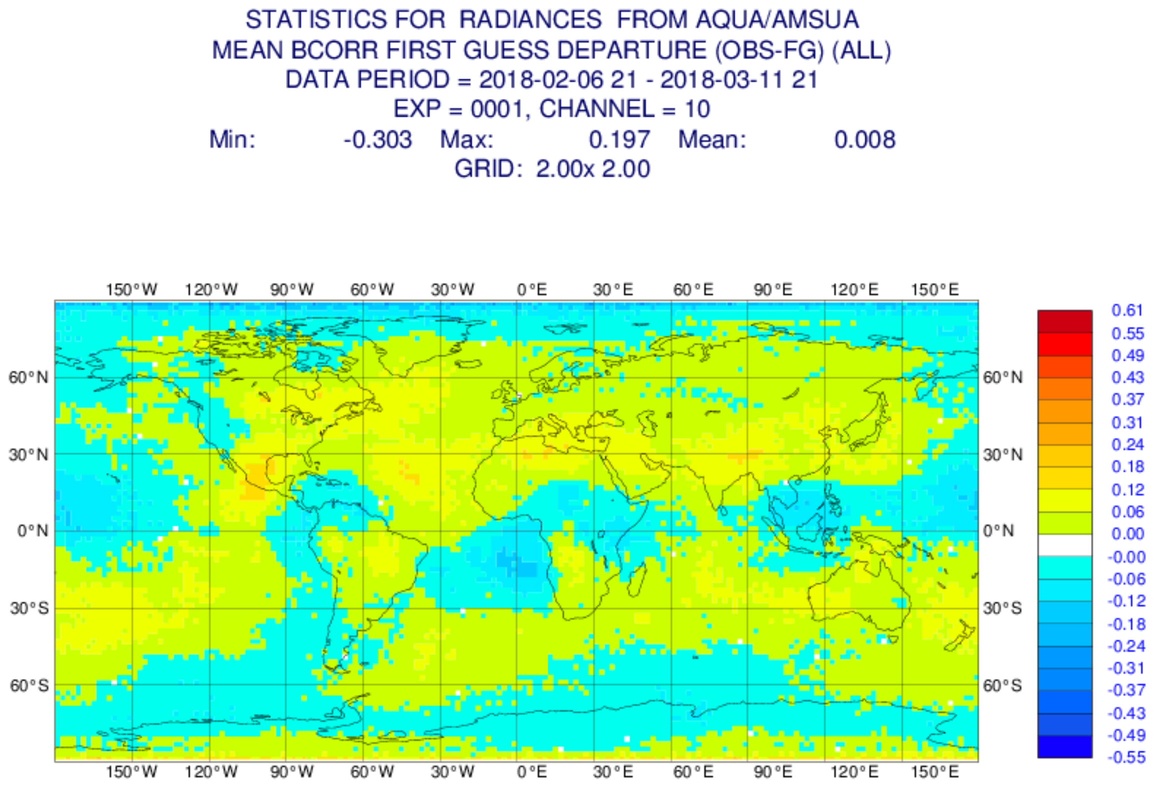
Mean background departures **before** bias correction for two similar channels on different satellites

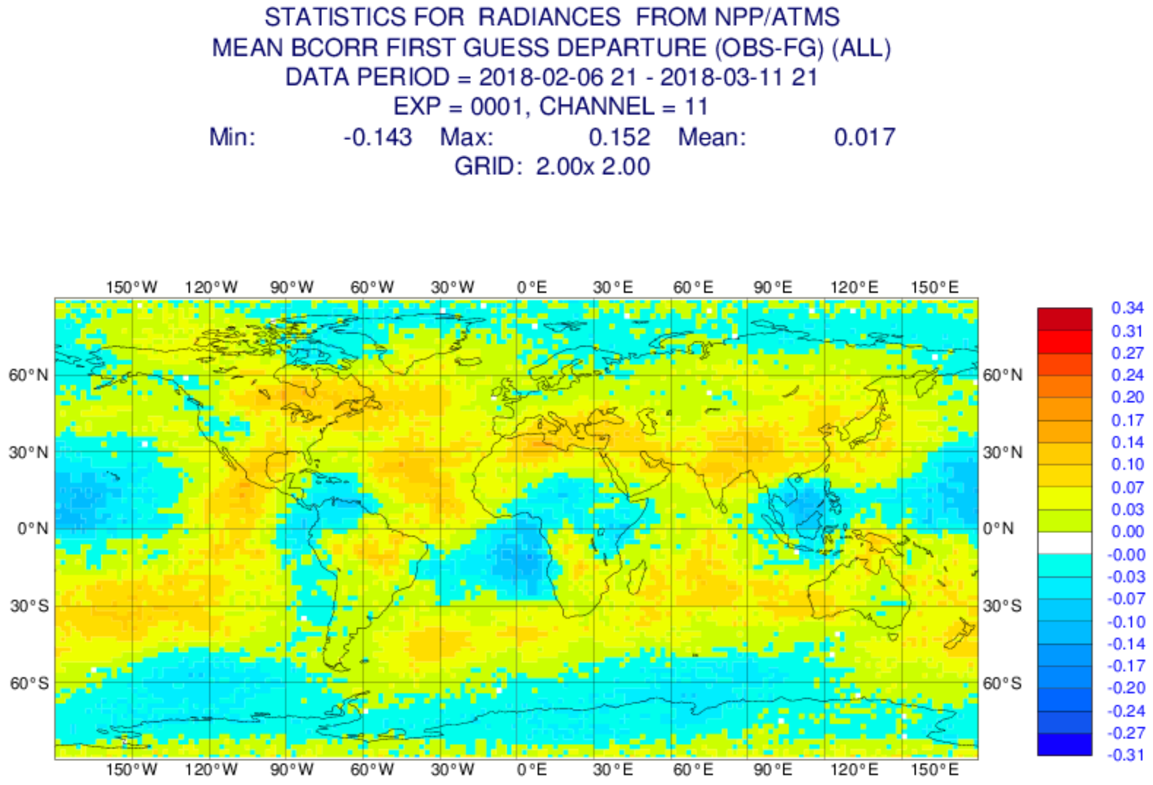




(note: different colour scales)

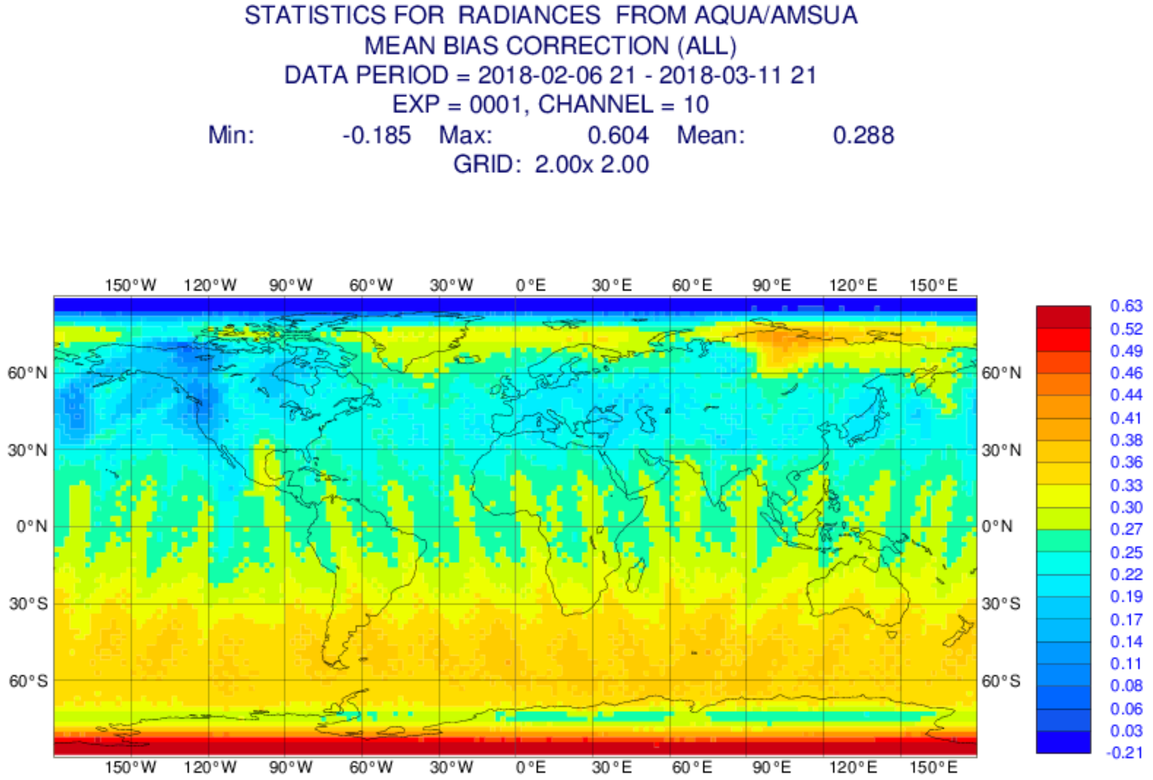
Mean background departures **after** bias correction for two similar channels on different satellites

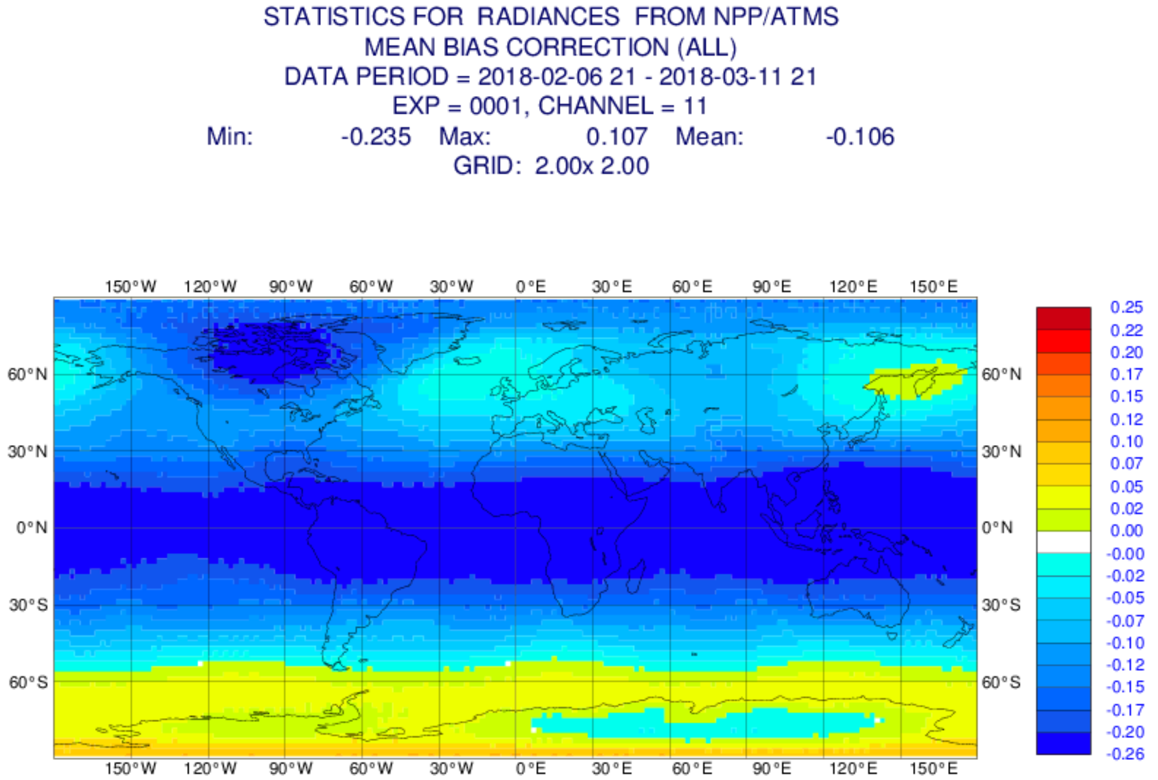




(note: different colour scales)

Mean **bias correction** for two similar channels on different satellites

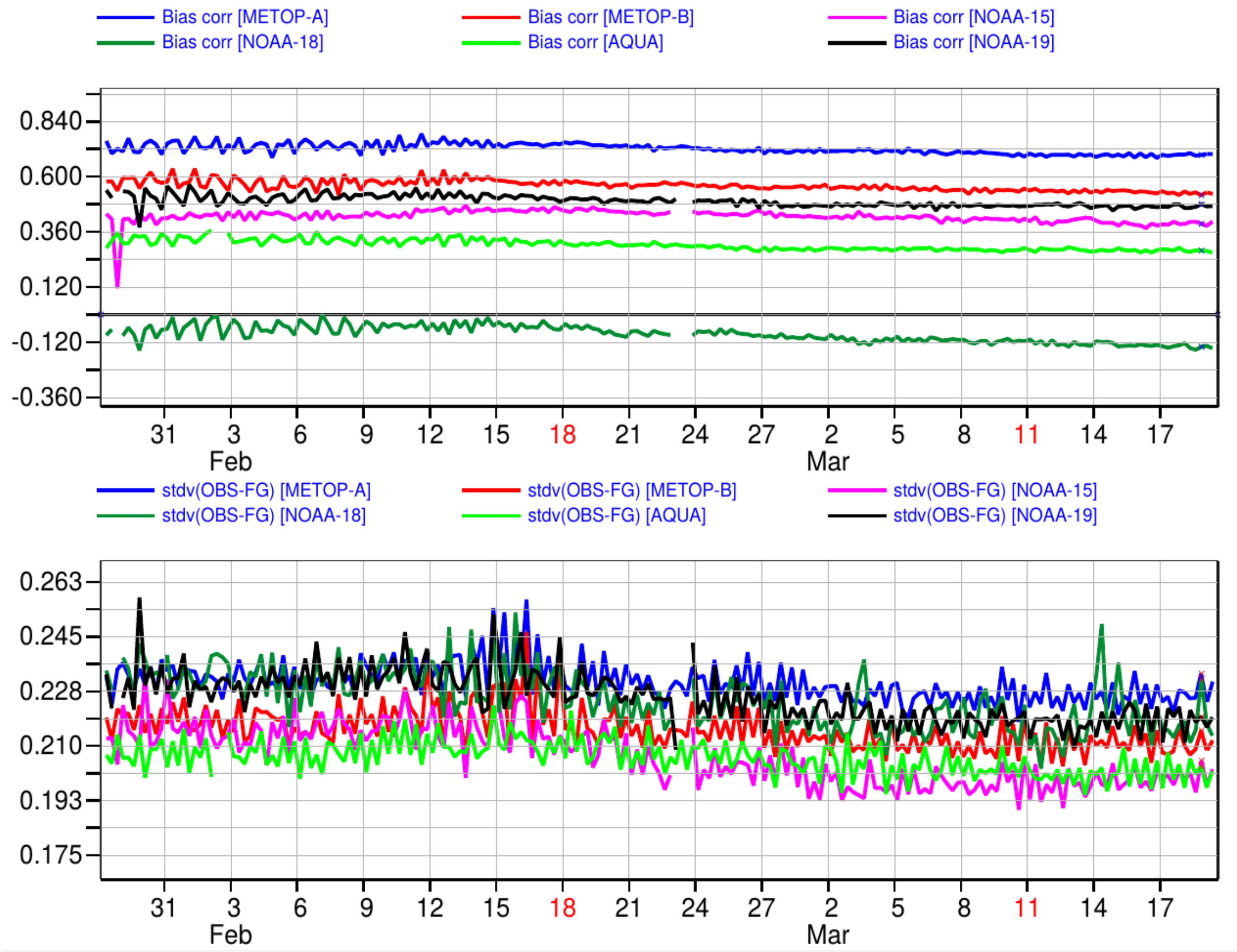




(note: different colour scales)

Time-series of departures for the same channel on different satellites

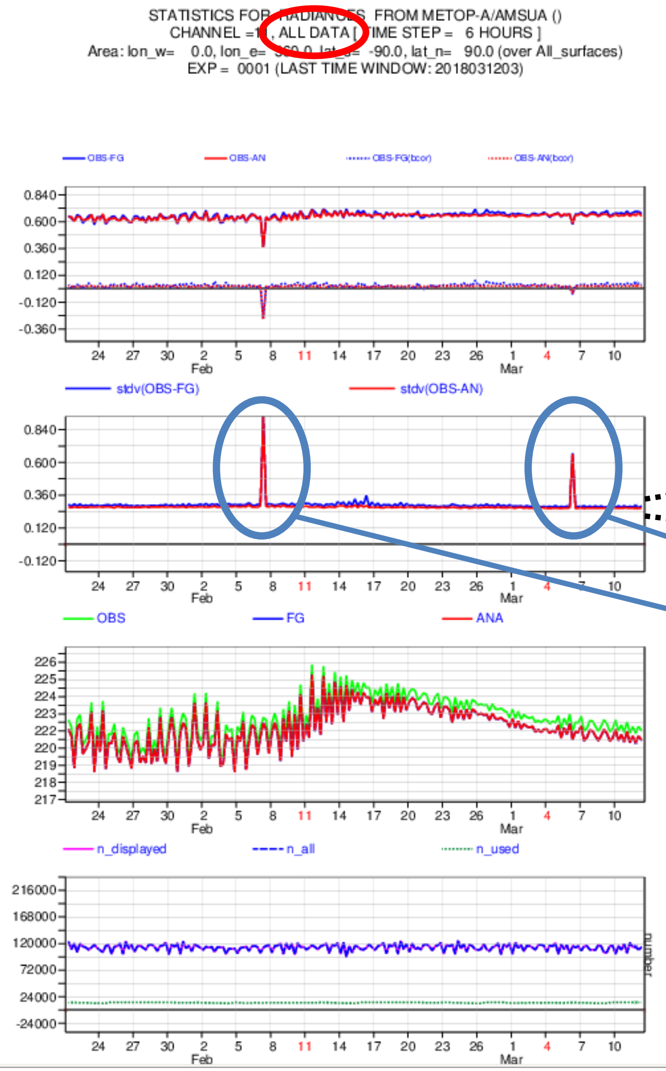
AMSU-A, channel 10, global statistics for used data



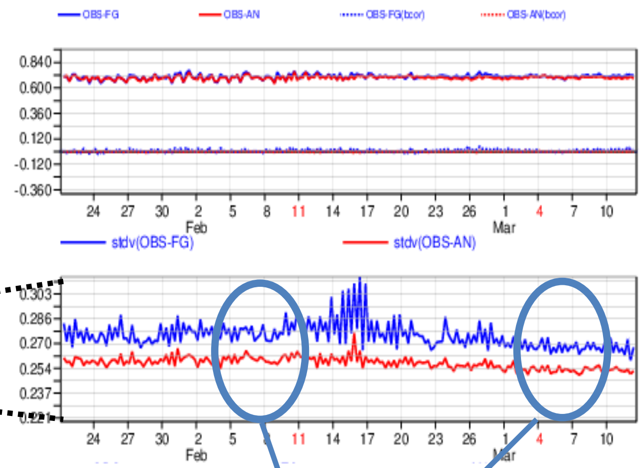
Bias correction

Stdev(o-b)

Departure statistics for different data selections

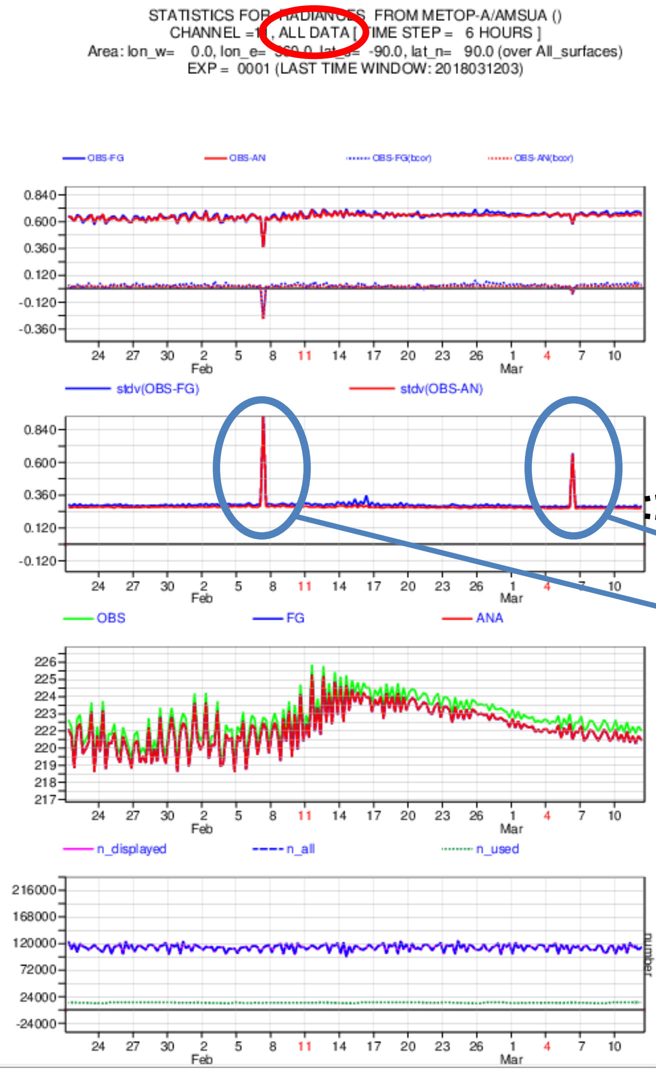


Compare with statistics for used data:

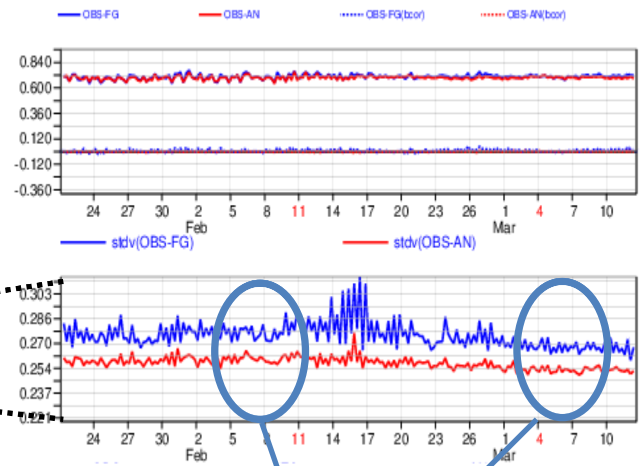


Poor data successfully removed by quality control in sample of used data

Departure statistics for different data selections



Compare with statistics for used data:

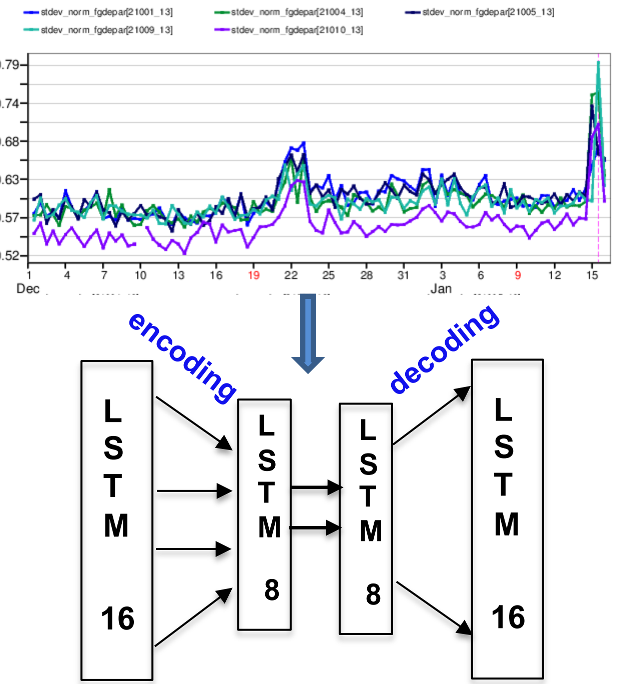


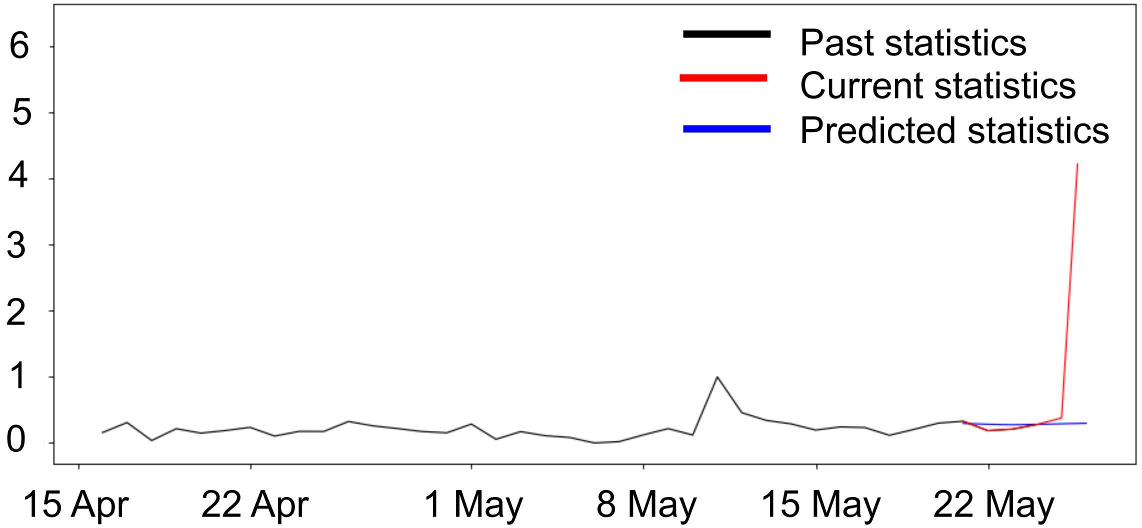
Poor data successfully removed by quality control in sample of used data

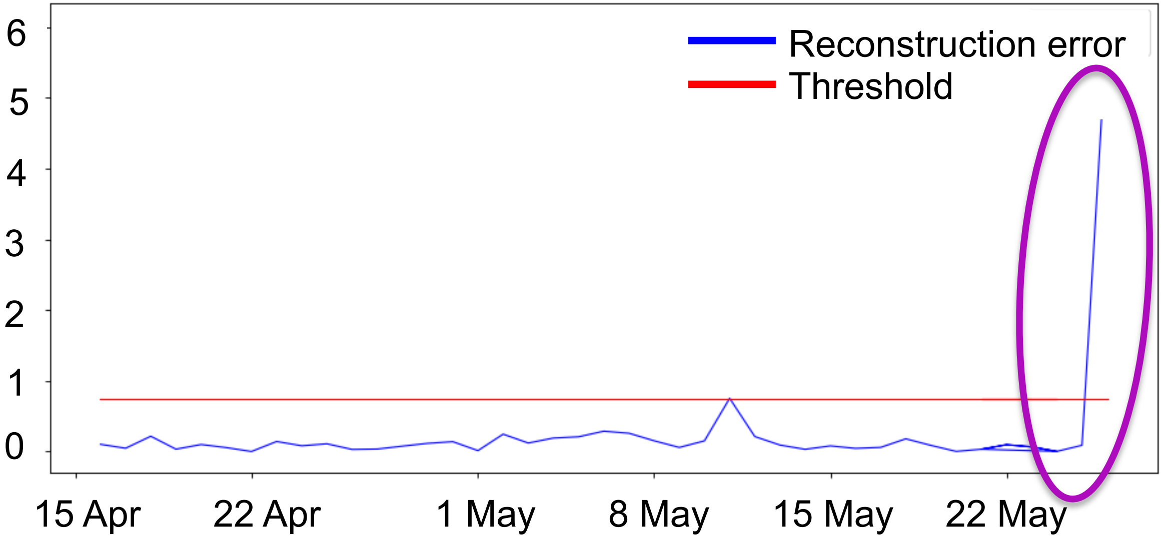
But what if the quality control is not sufficient?
With the large range of observations, manual checking for anomalies is not feasible.
<u>Need an automated system.</u>

Machine-learning based alarm system

- **Two-step concept:**
  1. **Detect anomalies** in time-series of observation statistics (e.g., stdev(o-b), mean(o-b), number of data)
     - Uses unsupervised learning, LSTM autoencoder (TensorFlow)
     - Compares actual time-series vs ML-predicted time-series constructed from past behaviour



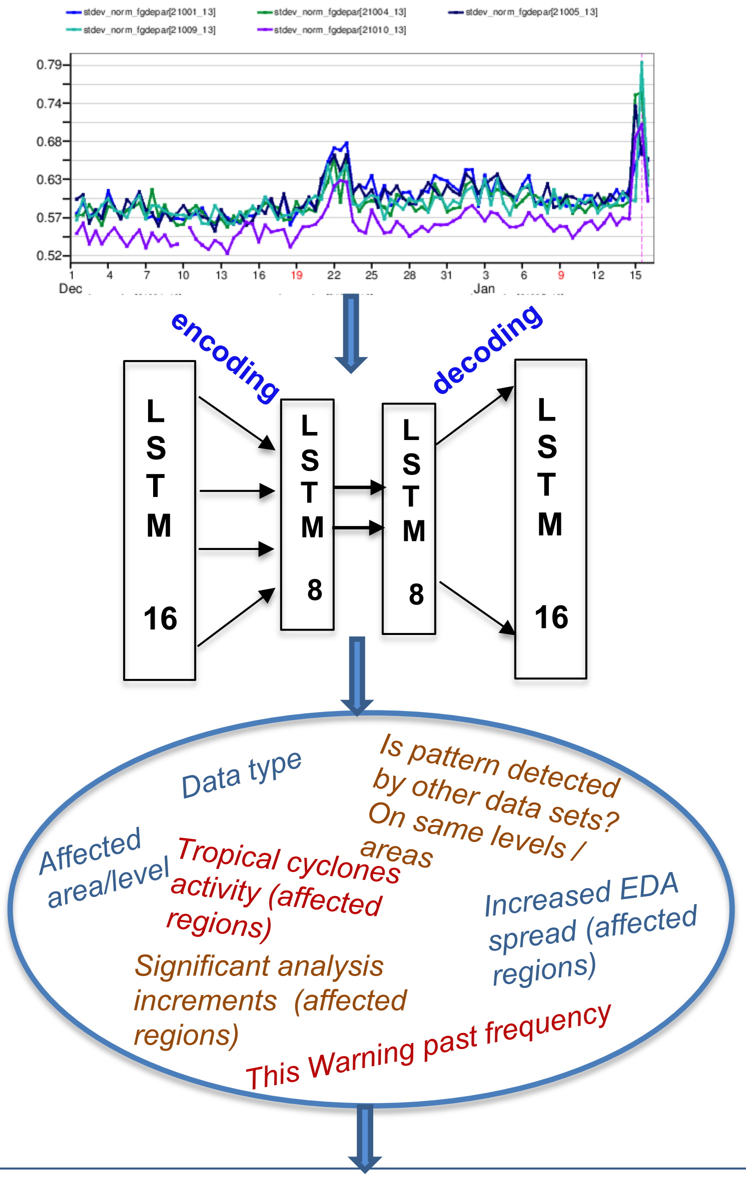




Machine-learning based alarm system

- **Two-step concept:**
  1. **Detect anomalies** in time-series of observation statistics (e.g., stdev(o-b), mean(o-b), number of data)
     - Uses unsupervised learning, LSTM autoencoder (TensorFlow)
     - Compares actual time-series vs ML-predicted time-series constructed from past behaviour
  2. **Classify anomalies**
     - Based on labelled warnings
     - Uses supervised learning, with anomalies from different instruments/channels combined with other auxiliary information as inputs
     - Random forest

- **Output:** Email warning indicating which data is affected, severity of anomaly and suspected reason



AMSUA

METOP-C AMSUA ALL SKY Radiances Global Channel 6: out of range:

(3 times in last 10 days)

2024022912_LWDA_0001_amsua1_21010_119_3_112_6.png

Severe: stdev_norm_fgdepar=0.91255, expected range:0.8221 0.8612

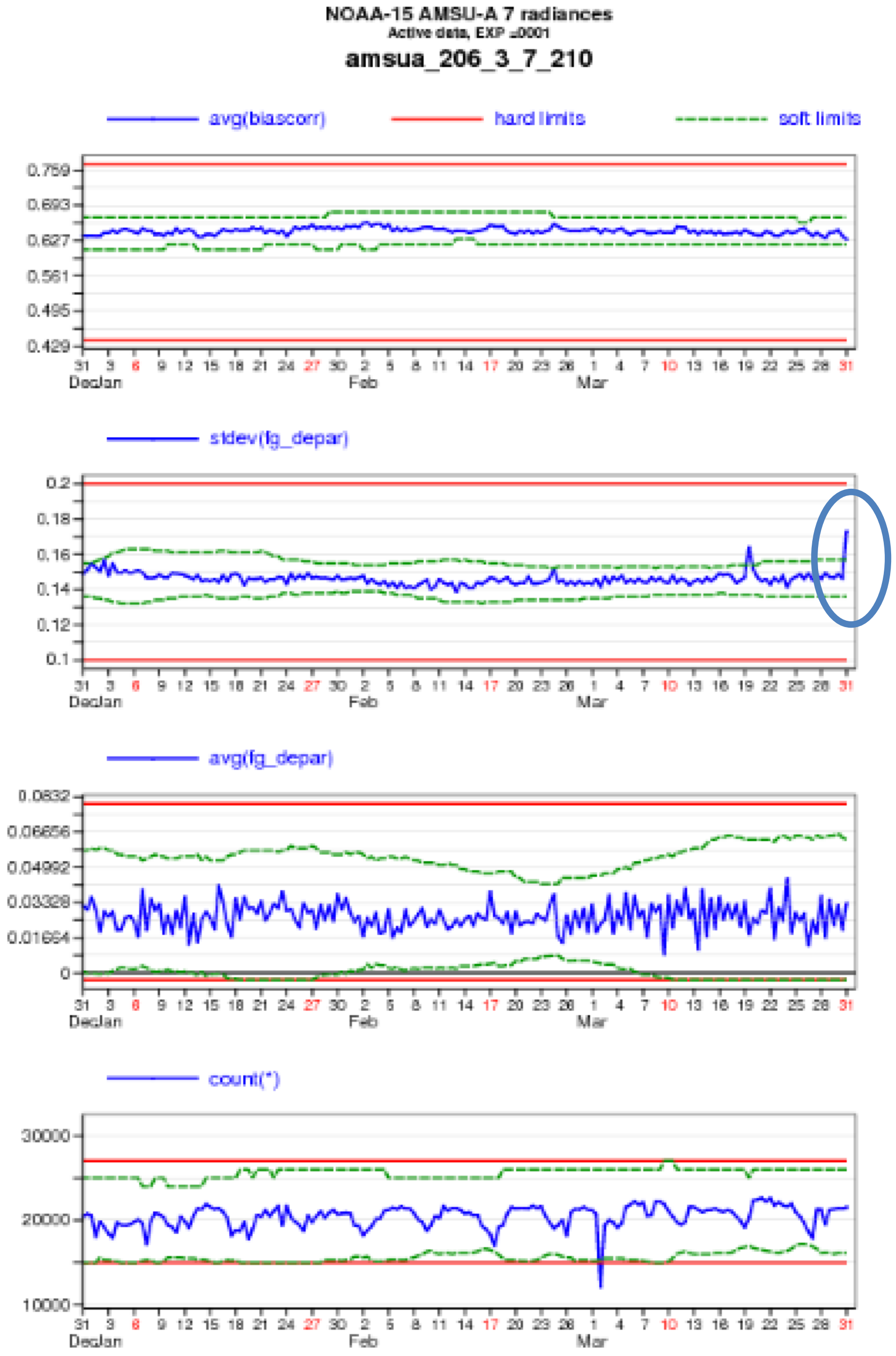
An example of an instrument noise problem…

AMSUA

NOAA-15 AMSU-A 7 radiances : out of range:

[amsua_206_3_7_210.png]

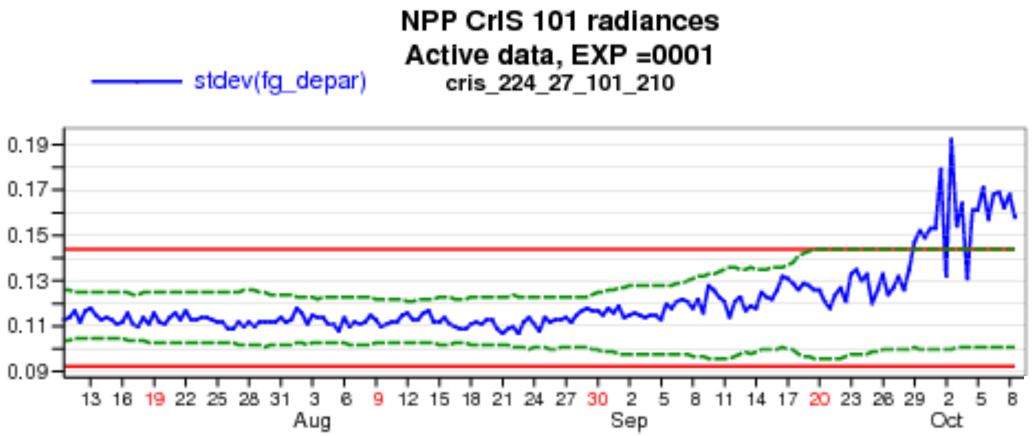
Severely: stdev(fg_depar)=0.172994, expected range: 0.136 0.157

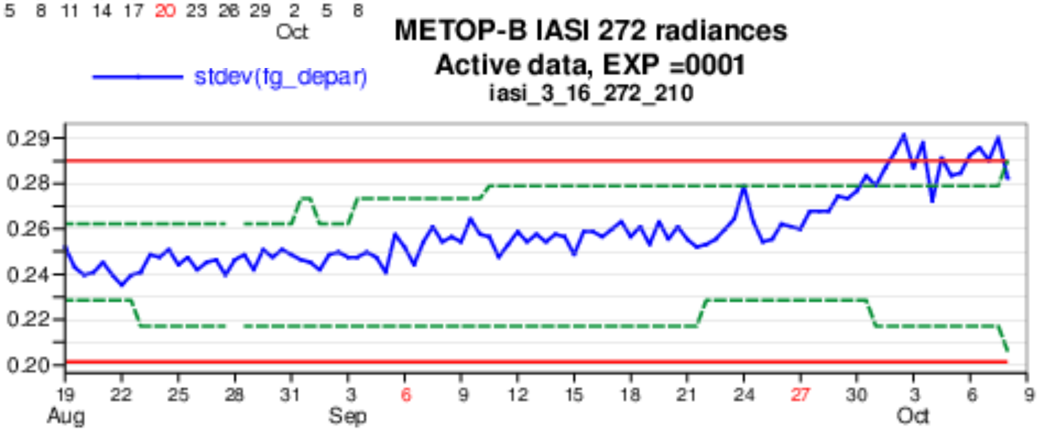


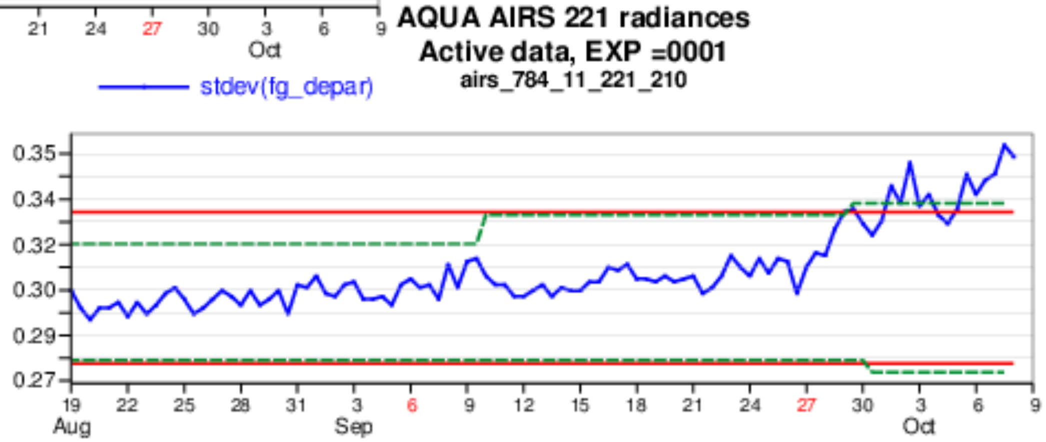
Whether or not to take action in such a case is a judgement call:

- It might be the beginning of the failure of the channel, so the channel should be excluded from assimilation as soon as possible.
- Or the problem might disappear tomorrow.

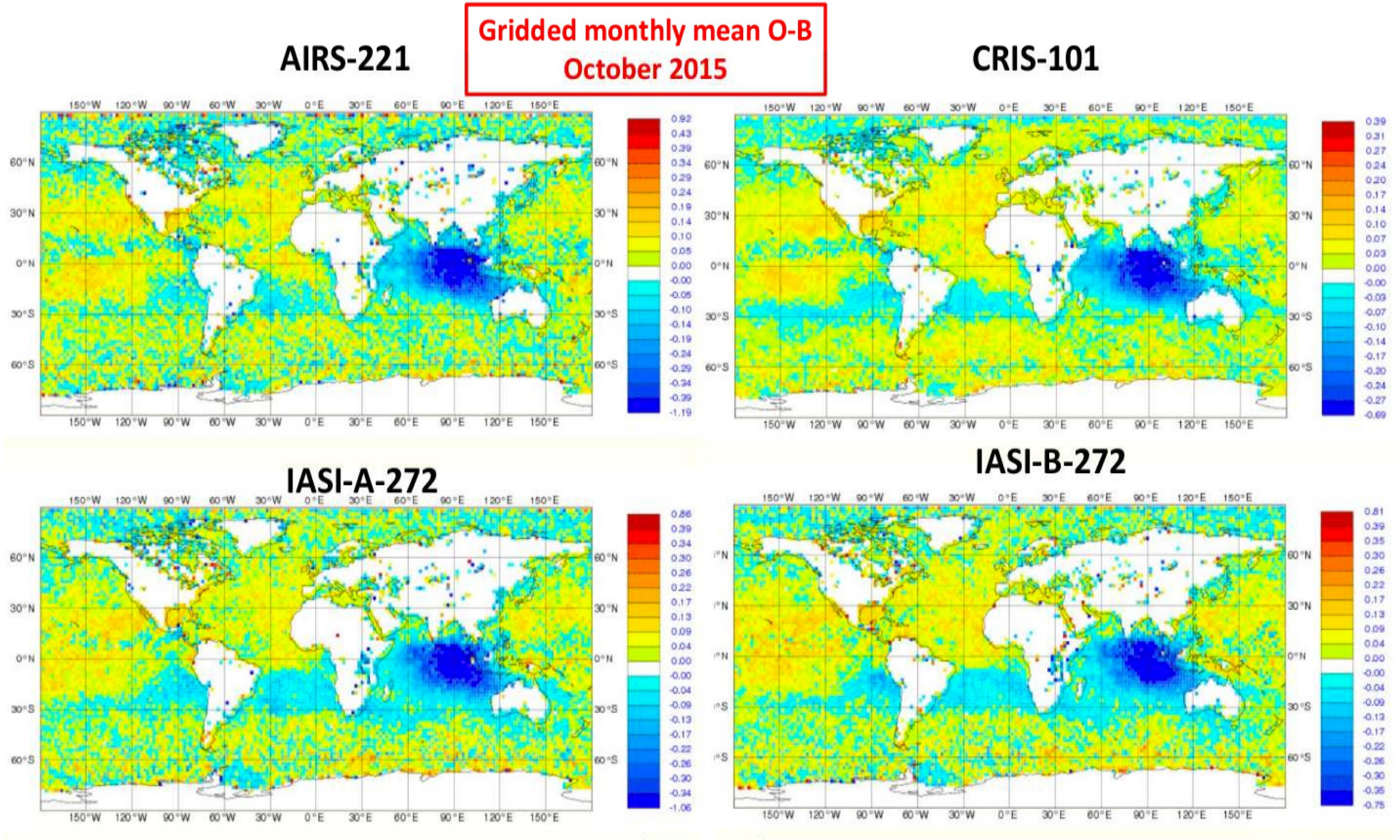
A different alert example…







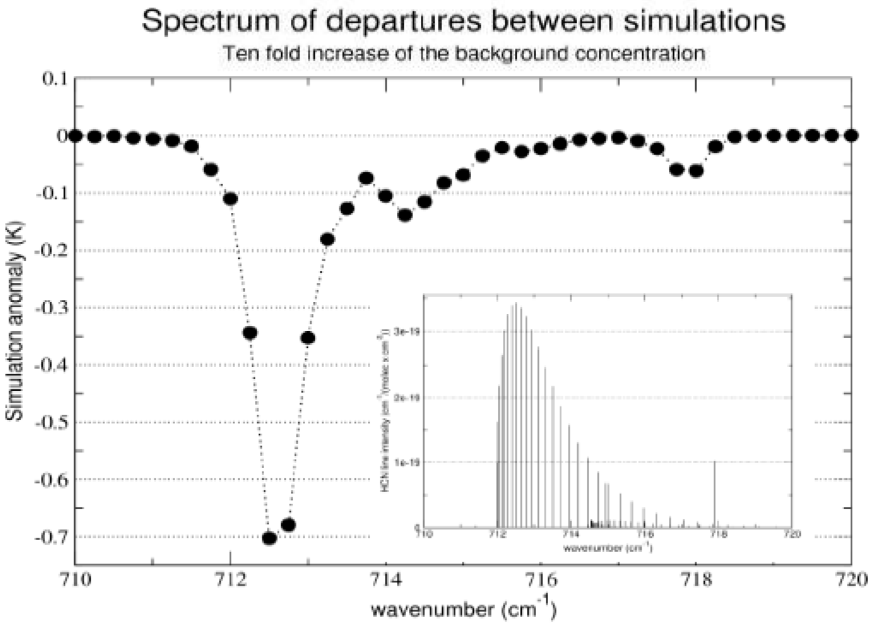
Alerts triggered by local bias signal in several IR channels around 712.5 cm^{-1}

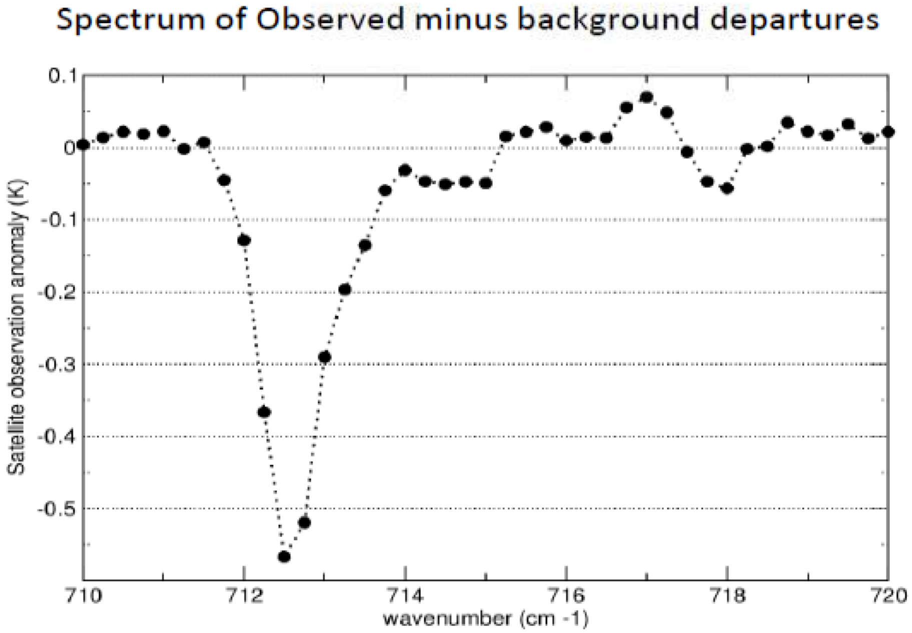


After some detective work….:
Local bias due to strongly increased levels of HCN over the Indian Ocean

Spectral signature of Hydrogen Cyanide (HCN)

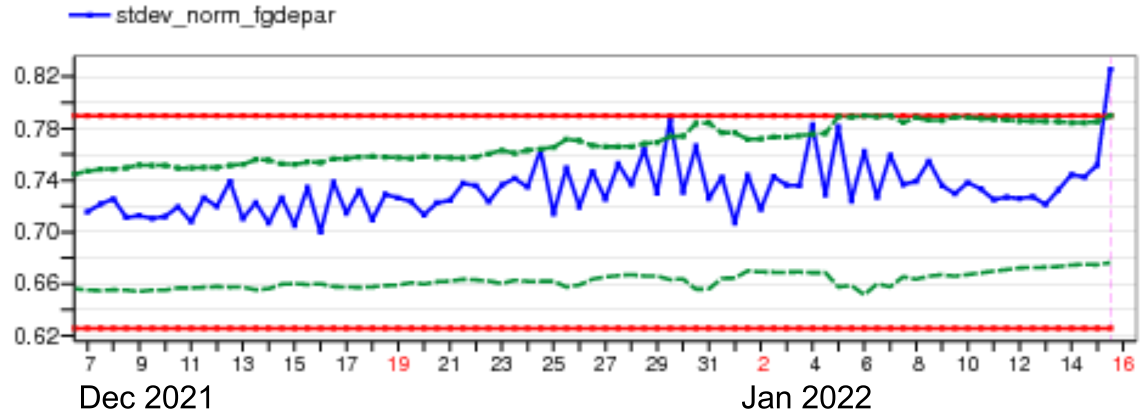
HCN is a known pollutant associated with biomass burning and the alarms coincided with the Indonesian fires



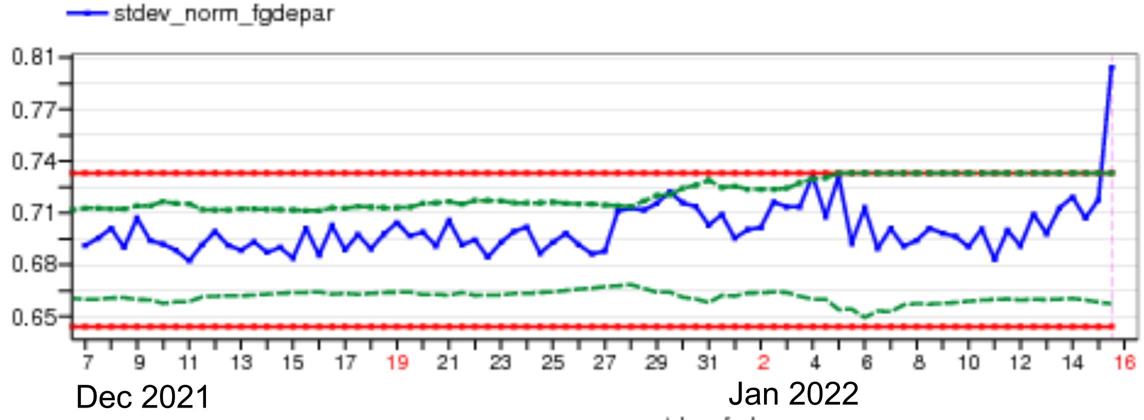


→ **Subsequently addressed through quality control**

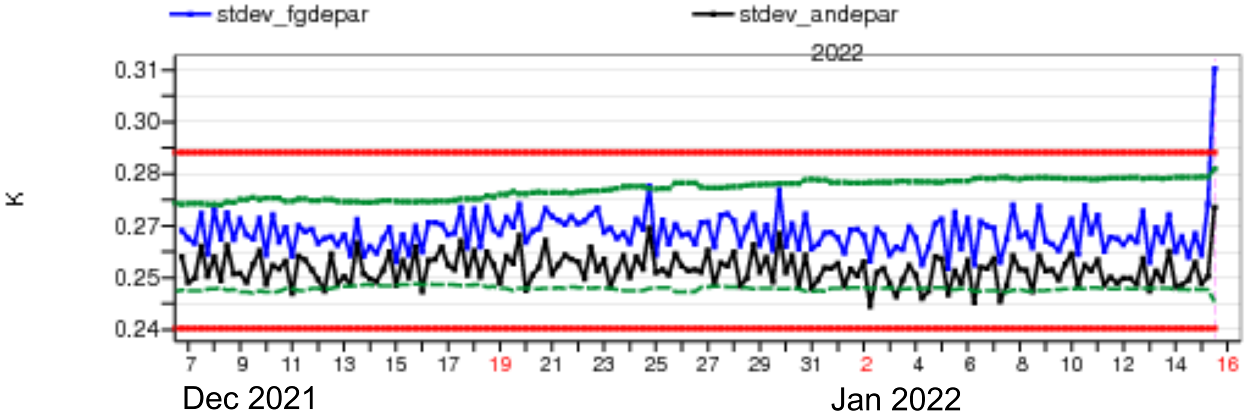
Another example… 15 January 2022



NOAA-18 AMSU-A ch 11



Metop-B AMSU-A ch 11

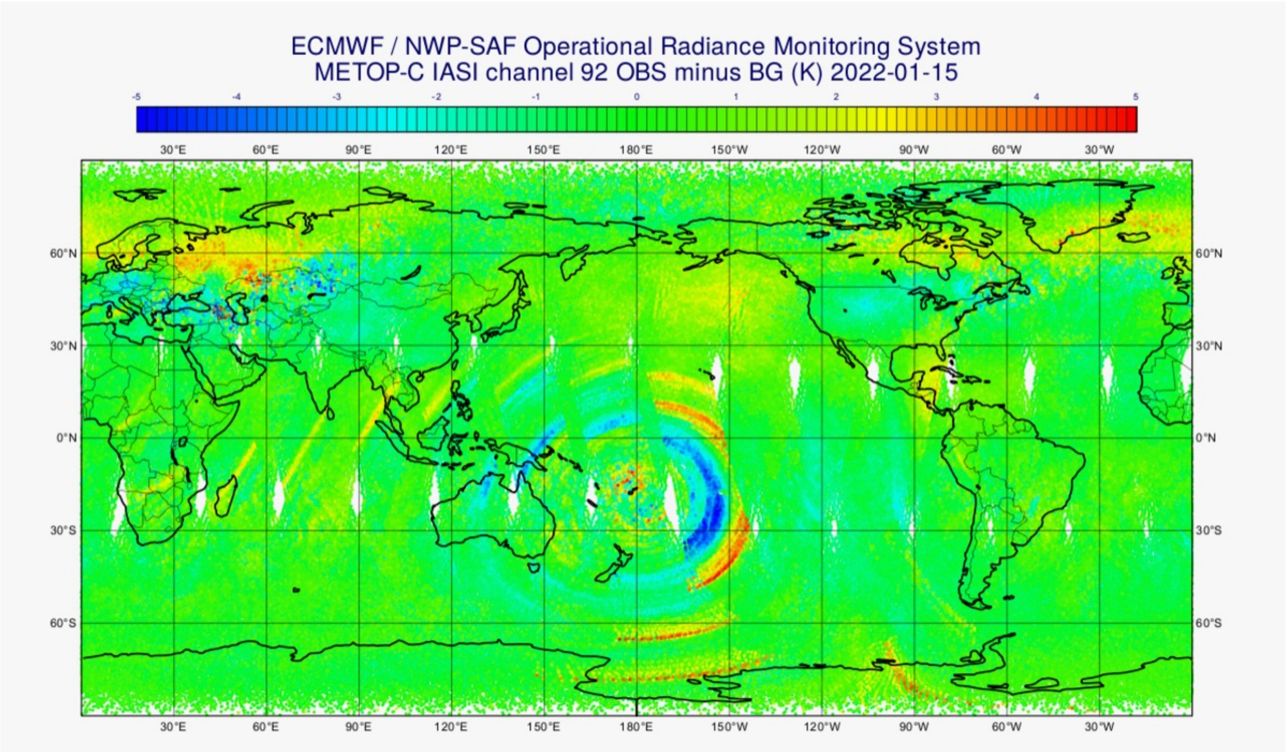


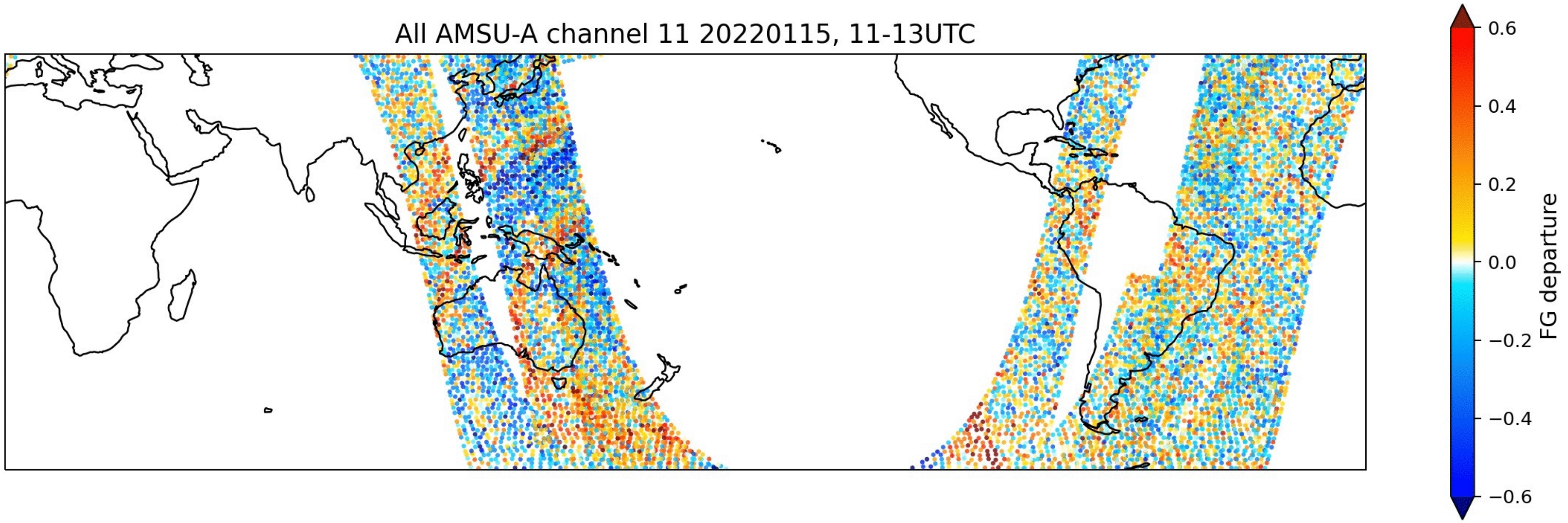
Metop-B
IASI ch 170

ECMWF

Shock-wave from the Tonga eruption

15 January 2022





ECMWF

Summary on data monitoring

- Monitoring of **departure statistics** is an essential aspect of data assimilation to
  - Diagnose “health” of the assimilation system
  - Diagnose model or observation biases
  - Characterise the quality of observations in the context of the wider observing system (→ contribution to satellite cal/val)
  - Characterise performance of bias correction schemes
  - Respond to sudden anomalies in observations
  - Etc.

Additional information

- Harris and Kelly, 2001: A satellite radiance-bias correction scheme for data assimilation. Q. J. R. Meteorol. Soc., 127, 1453-1468
- Derber and Wu, 1998: The use of TOVS cloud-cleared radiances in the NCEP SSI analysis system. Mon. Wea. Rev., 126, 2287-2299
- Dee, 2004: Variational bias correction of radiance data in the ECMWF system. Pp. 97-112 in Proceedings of the ECMWF workshop on assimilation of high spectral resolution sounders in NWP, 28 June-1 July 2004, Reading, UK
- Dee, 2005: Bias and data assimilation. Q. J. R. Meteorol. Soc., 131, 3323-3343
- Dee and Uppala, 2009: Variational bias correction of satellite radiance data in the ERA-Interim reanalysis. Q. J. R. Meteorol. Soc., 135, 1830-1841
- Han and Bormann, 2016: Constrained adaptive bias correction for satellite radiance assimilation in the ECMWF 4D-Var system. ECMWF Technical Memorandum 783.

Feel free to contact me with questions: Niels.Bormann@ecmwf.int